

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

**Defective images within this document are accurate representation of
The original documents submitted by the applicant.**

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SIANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 05:03:11 ; Search time 13643.2 Seconds
(without alignments)
862.151 Million cell updates/sec

Title: US-09-763-836-1

Perfect score: 713

Sequence: 1 gccagccctgatggggc.....atttgggtaaggtcatcgat 713

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.om.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.sy.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htgo.hum.*
- 31: em.htgo.inv.*
- 32: em.htgo.rod.*
- 33: em.htg.hum.*
- 34: em.htg.inv.*
- 35: em.htg.rod.*
- 36: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Match	Length	DB	ID	Description
1	695.4	97.5	9616	14	AB049088	AB049088 Hepatitis
2	692.2	97.1	9427	14	HPCJRNA	D14484 Hepatitis C
3	692.2	97.1	9616	14	AF139594	AF139594 Hepatitis
4	692.2	97.1	9379	14	AF139594	AF139594 Hepatitis
5	691.6	96.9	9418	14	HCV132996	AF207766 Hepatitis
6	690.6	96.9	1880	14	HCV132996	AF207766 Hepatitis
7	690.6	96.9	2540	6	HFC5TRJ4	D00832 Hepatitis C
8	690.6	96.9	2540	6	E04260	E04260 cDNA encodi
9	690.6	96.9	2540	6	E04805	E04805 cDNA to 5'
10	690.6	96.9	2540	6	E07391	E07391 cDNA encodi
11	690.6	96.9	9448	14	HPCJ483	D13558 Hepatitis C
12	690.6	96.9	9504	14	AB049094	AB049094 Hepatitis
13	690.6	96.9	9587	14	AF333324	AB049099 Hepatitis
14	690.6	96.9	9595	6	AR119832	AF333324 Hepatitis
15	690.6	96.9	9595	6	AR119832	Sequence
16	690.6	96.9	9595	6	AF054247	AF054247 Hepatitis
17	689.4	96.7	9533	14	AR119833	AR119833 Sequence
18	689	96.6	780	6	HCV132997	AJ32997 Hepatitis
19	689	96.6	9456	14	HPCRNA	AR095006 Sequence
20	689	96.6	9456	14	HPCRNA	D10934 Hepatitis C
21	689	96.6	9595	14	HPCJ491	D10750 Hepatitis C
22	688.8	96.6	9379	14	AF165059	AF054248 Hepatitis
23	688.8	96.6	9379	14	AF165059	AF165059 Hepatitis
24	687.4	96.4	2033	6	AF165060	AF165060 Hepatitis
25	687.4	96.4	2033	6	E08374	E08374 N-terminal
26	687.4	96.4	4987	6	E08869	E08869 5' region of
27	687.4	96.4	9431	14	HPCPO	E08872 DNA encodin
28	687.4	96.4	9431	14	HPCPO	D45172 Hepatitis C
29	687.4	96.4	9431	14	HPCPP	D30613 Hepatitis C
30	687.4	96.4	9471	6	E06261	E06261 cDNA encodi
31	687.4	96.4	9348	14	D89815	E06457 cDNA encodi
32	687.2	96.4	9379	14	AF165052	D89815 Hepatitis C
33	687.2	96.4	9400	14	HPCGENOM	AF165052 Hepatitis
34	685.8	96.2	923	6	AR153761	L02836 Hepatitis C
35	685.8	96.2	9335	14	D85516	AR153761 Sequence
36	685.6	96.2	9377	14	AF207763	D85516 Hepatitis C
37	685.6	96.2	9379	14	AF165051	AF207763 Hepatitis
38	685.6	96.2	9410	14	HPCKL12	AF165051 Hepatitis
39	684.8	96.0	9414	14	HPCUNKCDS	D50481 Hepatitis C
40	684.8	96.0	9415	14	HCU16362	M96362 Hepatitis C
41	684.2	96.0	9431	14	HCU45476	U16362 Hepatitis C
42	684.2	96.0	9600	14	AF176573	U45476 Hepatitis C
43	684.2	96.0	9605	14	HCU238799	AF176573 Hepatitis
44	684	95.9	9379	14	AF207758	AJ238799 Hepatitis
45	684	95.9	9379	14	AF207767	AF207758 Hepatitis
						AF207767 Hepatitis

ALIGNMENTS

RESULT 1	AB049088	9616 bp	RNA	VRL	05-DEC-2000
LOCUS	Hepatitis C virus genomic RNA, complete genome, isolate: HCVT094.				
DEFINITION	AB049088				
ACCESSION	AB049088.1				
VERSION	GI:11559442				
KEYWORDS	Hepatitis C virus (isolate: HCVT094, specific host: human) serum cDNA to genomic RNA.				
ORGANISM	Hepatitis C virus				
SOURCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.				
REFERENCE	1 (sites)				
AUTHORS	Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K., Hatanaka, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and Mishiro, S.				
TITLE	Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 9616)				

AUTHORS

Mishiro, S.
Direct Submission
Submitted (19-SEP-2000) to the DBJ/EMBL/GenBank databases. Shunji
Mishiro, Toshiba General Hospital, Department of Medical Sciences;
6-3-22 Higashi Oh-i, Shinagawa-ku, Tokyo 140-8522, Japan
(E-mail: shunji.mishiro@toshiba.co.jp, Tel: 81-3-3764-8981,
Fax: 81-3-3764-8992)

FEATURES

source

Location/Qualifiers

1..9616

/organism="Hepatitis C virus"

/isolate="HCVT094"

/specific_host="human"

/db_xref="taxon:11103"

/tissue_type="serum"

/note="HCV-genotype: 1b"

1..341

/note="complete"

342..9374

/codon_start=1

/product="polyprotein"

/protein_id="BAB18801.1"

/db_xref="GI:11559443"

/translation="MSTNPKPQKTKRNTNRRPQDVKFPGGQIVGGVYLLPRGPRRL
GVRAKTRTSQSRQROPIKARRPEGTWAGQYVNPPLYGNEGHWAGNLLSPRG
SRWSPGTPDRRRNRNLKVIDITLCGFADLMGYIPLVAGPLGGAARALAHGVLEED
SVNTATGNCPCVSIFELLALLSLCTTPASAYEVHNVSGIYHVDNCSNLSIVTEAD
MIMHAPGCVSRSDNARCVALTPTLAARNASVPTTIRRHVDLLVGAFAASMY
VGLDGSVFLISQFTSPRRHETVQDCNSIYPGHVTHGRMDMMNNNSPTTALVV
SOLLRIPOAVMDVAGAHGVLAGLAYVSMAGNAKVLIVMLFAGVDGNTHVSAGV
GRTVSSLANIFSLGASOKIOLINTNGSNHINRTALNCNDSLOTGFLAALYAHFNAS
CGERMASRISIDFDCWGPITVTAQSNLDORPCYHAPQQCGIIPASOCCGPTVC
FTSPVYVGTSDSGSPETYSWGENETDITLLNLRPPGNGWCTWNSGTFKTCGG
PPCNGVGNNTLCTDPCDFKHPKPEATYKCGSGPWLTPRLCLVDYPLRHHYPTVF
TIFKRYVGNTEHLHAACNWRGERCDLDRSELSPLLSTTEWQILPCSFETFL
PALSTGILHONIVDQVLYGIGSAVVFSAIKWEYVLLFLILLADARVACILWMMLL
IAOAEALENLVINAASVAGAHGFLSFLVFCAAWIKGRLPVGVAYALVGMVPLL
LLALPRAYAMDREMAASCGAVFVGLALLTSPYKYVLLARLIWLPFIAREAH
LOWVPLNVGRGDALILFCAARPELITDITLLALILGLPLVMFQAGIIRVIFYVR
AQLRVCLMVGRVAGHYVOMALMLAAATGTYYVDLTPLRDWAHAGLRDLAVFE
PVFSDMETKIIITWAGDAACDIIISGLPVSARRKEILIGPADSLRGRWRLLAPIT
AYSQOTRGLGCIITSLTRGDKQVEGEVQVSTATQSFASVNGVCWTVYHGASGK
TLAGPKPIQMYNVLKSDGGLPLCPSHVVGIPRAACTRGVAKAVDFVPVSEMT
MRSPVFTDNSPPAVPQTFQVAHLHAPITGSGKSTKVPAAAYAAQYKVLVLPVQ
GDAYMSKAGHTDPIRIGIKITITGAPITSTYTGKFLADGCGSGGAYDIIICDEHS
TGSTIILIGTVDQAEAGARLVLTATPPGVSIVPHNIEVALNSITGEIPFYK
AIPTEALKGRHLIFCHSKKDELAALSLGINATYRGLDVSIVPTSGDVVVA
TDALMTGTFGDFSDVDCNTCTVDFSLDPTFTIETTVPQDAVSRTOGRGTGRG
RRGIYREVTPGERPSGMDSSVLCGYDAGCAWELTPAETSVRRLAYLTPGLPVQ
DLHFEVESFTGLTHIDAHFLSQTKQAGDNFPYLVAYQTVCAARAPPFSWQGMKC
LIRUKPLHPTPLULYRLGAVHNEVTLHPVTKYIMACSADLEVSTVWLVGVLA
ALAAAYCLTGSVIVGRIILSGRPVAPIDREVLREFDEMECASHPLYIEQGNQALAE
QFKQKALGLQATVQAEAAAPVVEKSRALAEFAWAKHMNFISGIIYLAGSLTPGN
PATASLMFTAITSPLTOSTLFLNLTGWWAAQAPPSAASAFVAGLAVGAGVSI
GLKGLVLDIAGYAGVAGALVAFKVMGMSPEDEIVNLLPATLSFCALVGVCA
ILRHVPGGAGVQVMNRLTAFASRGNHVSFTHVVPESDAAARVQLSSLTIIQLLK
RLHWINEDSCPCSGWLRDWDWICTVLTDFKTLQSKLLPLPLPFPFSQGRYK
GWRMGDMGIMTQPCGAQISGHVKNVSMRIVGPKTCSNTWHGFTPIINAYTGTPTSP
APNYSRALRVAAEEYEVTVRGVDFHYVTGTMTDNLKPCQVPTPEFTFEDVGRVLRH
YAPACKLLREEVVFOVGLNOYLVGVSOLPCEPDVAULTSMLTDPDSHITAEAKRL
ARGSPPLSSASSASQSPSLKATCTTHDSPOADLEALNLMWQMGNI SRVSEN
KVIILDSFDPLRAEDEREVSVAELILKRSREPPALPIWARDYNPLPLESKWDPIY
VPPVHGCPLPTKPPPIPPRRKRTVLTSTVSSALAEATKTFSSGSSVDSGT
ATAPDPTSDGQSDVESYSPNPLEGPDPLSDGSNSTVSEAGEDVYVCCNS
YTWTGALITPCAAEESKLPINALSNLRHNMITYATSRAGLRKQKFTFDRLOVLD
DHYRDVLEKMAKASTVKALKVEEACKLTPPHSAKSKFYGAKGVNRLSSKAVNH
RVSKMDLLEDSETPIDTTIMAKNEIFCVQPEKGRKPARLIVPDLGVRCVEMALYD
VWSTLPOAVWSSYGFQYSPQCRVEFLVNTWKKCPMGFSYDTRCFDSIVTENDIRI
EESIQCCLAPARQARISLTERLYVGGPLTNSKQNGCYRRKRSAGVLTSCGNPL
TCYLKASCAARAKLQDCTMLVNGDDLVICESAGTOEDAASTRATEMYSTRASPG
DTPQPEYDELITSCSSNSVAHDAGSKRYLYLRTDPTPLARAAMEATHTPWNWL
GNITIMAYPTLWARMILMTFFSILLAQOLEKALDQOYGCYSIEPLDLPQITELRH
GLSAFSLHSVSPGEINRVASCLRLKLVPLLRVWHRHRSVRKLLSOGGAAATCGKYL

FNMAVKTKLKLTPIPAASRRLDSQWVAGYSGGDIYHLSRARPRWFMLCLLLLSVGV
GIYLLPNR"
9375..9412
/note="upstream of poly-U stretch"
9413..9493
/note="poly(U) length is indistinct
poly(U) stretch"
9494..9616
/note="3' X tail"
BASE COUNT 1920 a 2861 c 2702 g 2131 t 2 others
ORIGIN

Query Match 97.5%; Score 695.4; DB 14; Length 9616;
Best Local Similarity 98.5%; Pred. No. 1.3e-167;
Matches 702; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 gccagccctgatgggggacactccaccatagatcaactccctgtgagaaactatg 60
|||||
Db 1 GCCAGCCCGGATTTGGGGGACACTCCACCATAGTCACTCCCTGTGAGGAACACTG 60
Qy 51 tcttcacagaaagcgtctagccatggcgttagtatagatgctgcagcctccaggac 120
|||||
Db 61 TCTTCACCGCAGAAACCGTCTAGCCATGGCGTTAGTATGATGCTGTCGAGCCTCCAGGAC 120
Qy 121 cccctcccgaggagagccatagtgctgctgcggaaccggtgagtacacgggaattgcag 180
|||||
Db 121 CCCCCCTCCGGGAGAGCCATAGTGTCTCGGGAACCGGTGAGTACACCGGAATTCGCAG 180
Qy 181 gacgaccgggtctcttcttgatcaaacccgcctcaatgcctggagattggggtgctccc 240
|||||
Db 181 GACGACCCGGTCTTCTTGTGATCAACCCGCTCAATGCTGAGATTTGGCGTGCCTCC 240
Qy 241 gcgagactgtagccagtagtgctgctgcggaaccggtgagtacacgggaattgcag 300
|||||
Db 241 CGGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGGCCCTTGTGGTACTGCTCATAGG 300
Qy 301 gtgctgtagtgcctcccgagggtctctgtagaccgtgaccatgagcacaatccctaaac 360
|||||
Db 301 GTGCTTGGAGTGCCTCCCGGGAGGTCTCTGTAGACCGTGCATATGACGACAAATCTCAAC 360
Qy 351 ctcaagaaaaaacaaagtaaacaccacccgcccacagagatcaagttcccgagcg 420
|||||
Db 351 CTCAAAGAAAAACAAACGTAACACCAACCCGCCGCCACAGGACGTTAAGTTCCCGGCG 420
Qy 421 gtggtcauatcgttggtgagtttacctgttgccgcgcaggggcccaggttggtgtgc 480
|||||
Db 421 GTGCTCAGATCGTTGGTGGAGTTTACCTGTTCGCCGCGAGGGGCCCGAGGTTGGGTGTC 480
Qy 431 gcgcgactaggaagactctcccgagcgtgcgcaacctcgtggaaggcgacacactatcccca 540
|||||
Db 431 GCGGACTTAGGAAGACTTCCGAGCGGTGCGCAACCTCGTGTGAAGGCGACAACTATCCCA 540
Qy 541 aggcctcccgcccgagggcagaccctgggtgcagccgggtatctctggccctctatg 600
|||||
Db 541 AGGCTCGCGCGCCGAGGGCAGGACCTGGGTTCAGCCCGGGTATCTCTTGGCCCTCTACG 600
Qy 601 gcaacgagggcatggggcgagcagatgctccctgctgcgcgcgcccgcgclcccgccctagt 660
|||||
Db 601 GCAACGAGGCGATGGGTGGCGAGGATGGCTCTCTGTACCCCGCGGTGTTCCCGGCGCTAGT 660
Qy 651 ggggccccttcggaccccccgcgtaggtgcgcgtaatttgggtaaggtcatcgat 713
|||||
Db 651 GGGGCCCCACGACCCCGCGGTAGGTGCGCTAAATTTGGGTAAAGTATCATCATGAT 713

RESULT 2

LOCUS HPCJRNA 9427 bp RNA VRL 03-FEB-1999
DEFINITION Hepatitis C virus strain J33 genomic RNA, complete genome.
ACCESSION D14484 D01173
VERSION D14484.1 GI:1160327
KEYWORDS complete RNA genome.

KEYWORDS

SOURCE Hepatitis C virus.

ORGANISM

Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE

1 (bases 1 to 9616)

AUTHORS

Beard, M.R., Abell, G., Honda, M., Carroll, A., Gartland, M., Clarke, B., Suzuki, K., Lanford, R., Sangar, D.V. and Lemon, S.M.

TITLE

An infectious molecular clone of a Japanese genotype 1b hepatitis C virus

JOURNAL

Hepatology 30 (1), 316-324 (1999)

MEDLINE

95315771

PUBMED

10385673

REFERENCE

2 (bases 1 to 9616)

AUTHORS

Beard, M.R., Abell, G., Honda, M., Carroll, A., Gartland, M., Clarke, B., Suzuki, K., Lanford, R., Sangar, D.V. and Lemon, S.M.

TITLE

Direct Submission

JOURNAL

Submitted (30-MAR-1999)

JOURNAL

Microbiology and Immunology, The University of Texas Medical Branch at Galveston, 301 University Blvd., Galveston, TX 77555-1019, USA

FEATURES

Location/Qualifiers

1. 9616

/organism="Hepatitis C virus"

/strain="HCV-N"

/db_xref="taxon:11103"

/note="infectious full-length clone"

genotype: 1b"

342..9389

/note="encodes structural and non-structural proteins"

/codon_start=1

/product="polyprotein"

/protein_id="A044718.1"

/db_xref="GI:5532422"

/translation="MSTNPKRQKTKRNRPPQDKVFGGGQIVGGVYLLPRRGRPL

SRVAKTSRSRQRRQPKPKARRPEGRMAQGYPPWLYGNEGMAWGLVLLSPG

GSPWGTGTPRRSRNLGKVIDTLCGLADLMGYIPLVGGPLGGAARALAGVRLVD

GVNATNLPCGCSFIELLALISCLTPASNEHVRNASGVYHYVNDSCNSIPEAD

LIMHTPCVPCVRGNSRCWALTPPLAARNATIPITIRHVDLLVGAALCSAMY

VGLDGSVYLSOLFSPRRHATLQDNCISYPGHASGHRMADMMNNSPTALVY

SQLRIPQAVIDWAGHVGVLAYISMGNNKAVLIIVMLFAGVDHGLTGHGA

ARLTSGFAGLTGPGSRIQILINTGSHWNRNALNCNDSLOTGFLAALFYAHFNS

CPRMASCSDIKFDQGWGPIVABPTKDPDRPCWYHYPPOCGVLPASOVCGPVY

CFTEPVSVDTRGLNPTYSWGENPTDVLNANTRPQGNWCTWNSSTGFTKCG

APFCNIGVGNLTCTDCEFRKHEPATYSKCGSPMLTPCWVDYPIRLWHTPTCVN

FSIPKMYVGGVHRUNAACNTRGRNCNDDDRSLSPLSTTEWQVLPSCFTL

LPALSTGLILHQINVDVQYLYGTSVAVFAIKWEYVLLFLLLADARVACLMML

LIAQEAALNVALNAASVAGHGLISFLVFFCAWYIKRGLVPGAAYAFYGAWPL

LLLTLPPEAYAMDREAASGGAFFVGLALLTSPYKYVFLARLLWLOVLTTRA

HLHWVPLNVRGRDAIILLTCAVHPELIFDTKLIALILGLPLMWLOAGITRPVYV

RAOGLIRACMLVRKAGGVYQVAFMFGALTYVYNHLTPLRDWAHAGRLDIAV

EPVFSDMETKIITWADTAACGDIILGLPVSARRREIILGLPADSLVGGWRLLAPI

TAYSQTRGLGCIITSLTRDNQVVEGVYSTATQSFATCVNGVCWTVYHGAS

KTLAAGPKPTAQYNTVDQLVGPAPPGARSPTCTCGSSDLYLVTRHADYVPRR

DNRGSLSPRPVSYLYKSSGGPLLCPSGHAVGVFRAAVCTRGVAKAVDPVPMET

TMRSPVTDNSPPAVPQTFQVAHLHAPTSGKSTRVPAAYAAQGYKVLVNPVAT

LGFGAMSKAHCIDPNTRTGRTITCAPITYSTYKFLADGCGSGAYDILMDECH

LSTSTTLGLGLVDQETAGRLVATAPPGSVTVPHNIEEVALSNTEIPFGY

KATPIEAKGRRLIFCHSKKCDLAAKLSGLGINAVYIRGLDVSIVPTSGDVVY

ATDALMTYTGDFDSVDICNTCTVDFSLDPTFTTETTPQDVAISRQRRTGR

GRGCIYRVFTGPRSPGMSFVLCYDAGCAWYELTPAETSVRLAYRNTPLCPVC

ODHLERWESVTGITHDAFLSQTQAGDNFYLVAQYATCARAQAPPSNDQWK

CLIRLAKTFLHGPPLLYRLGAVQNEVTLTHPTIKYIMACMSADLEVVTWVLSLITQ

AALAACTLTGVSIVVIGRIILSRGPVAVPREVLYREFDEMECAASHLYIEQGMOLA

EOPKALGLLOTTAQEAAPVVEKSKRALETFWAKHKNFISGTQYLAGSLTLP

NPATLSLMTASITSLPTNTLLNFGLVAAOLAPPASAFAVGAIAAGVS

IGLKVLDILAGVAGVACALVAFKVMGSEAPSAEDLIINLLPAILSPGLVGVVCA

AILLRHVHPEGEGVQWNRLLAFASRCNHNHSPHYVPESDAARVQVLSLITQ

KRLHOWINEDCTPSCSWLURDWMVCTVLSDFKWLQSKLLPLRPGLPVFLSCQY

KGVWRGDIMHTPCPGAGIAGHVKNGSMRIIGPKTCNTWGHGTFINAYTGPCTPS

PAPYKALRWAAEEYEVYEDHVTGIIITDNVQPCQVPAPPEFFTEVDGVRILH

PAPVCKPPLDVRVQGLVYLSQLPCEPEPDVAVLTSMLTDPSSHVTAAAKRR

LARGSPSLASSASOLAPSLRATCTTHSSYNLSDPDVLLIANLLWRQEMGNTR

KNPDYVPPVHVGCPPLVPKAPP IPPRRKRITVTLDTSTVSVLAELATKTFGSSELSA
ADCSMTATAPDQTSNDGSKDSADSCSMPLLEGEPDQDPSLDSGWSVTSSEAGESVY
CCSMTYWTGALITPCAEESEKLPINALSNLRLHNNVATTSRSAGLRQKKTFTDR
LOVLDDHVRVWLKEMKASTVKAKLLISEACPLTPHSAKSKFGYCAKDYRLNLSR
AINHRSVWEDLLEDTYPTDITVMAKNVFCVQPEKGRKPARLIIVFPDUGRVCEK
MALYDVVSTILPQAVMGSSYGFYSPGQVRBFVLKAWKSNKPMGFSDYTRCFDSTVTE
NDRIVESIYQCCDLAPAKIAKSLTERLYIGPLTNSKQSCGYRRCRASGVLTTS
CGNFLTCLYKASACRAAKIQDCTMLVNGDVLVVICESAGTQEDAASLRVTEAMTRY
SAPPGDLPOPEYDELELITSSSNVSAHDSGKRVYLLTRDPTPIARAAWETARHTP
VNSWGLNIMVAPTLMWMLMTHFFSILLAOEOLKALOCQIYAGCSTEPDLDPQI
IERLHGLSASLSHSYSGEINRVASCLKGLVPPLRVWRURADVRKAKLLSQGGAAT
CGKYLFWNAWTKLKLTPIPAASOLDLSGMFVAGYSGGDIYHSLSRARPRWFMCLCLL
LSVGVGYLLLPNR*

HASE COUNT 1891 a 2892 c 2734 g 2099 t
ORIGIN:

Query Match 97.1%; Score 592.2; DB 14; Length 9616;
Best Local Similarity 98.2%; Pred. No. 8.3e-167;
Matches 700; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 gccagcccccgtatggggggacactccacatagatcactccccctgtgaggaactacg 60
Dy 1 GCCAGCCCCCGATGGGGGGCACACTCCACATAGATCACCCTGTGAGGAACACTG 60
Qy 51 tcttcacgcagaaagcgtctagccatggcgttagtagtgctgagtcagcctccagac 120
Dy 51 TCTTCACGCAGAAAGCGCTAGCCATGGCGTGTAGTATGATGATGCTGCGAGCCTCCAGGAC 120
Qy 121 ccccccctccgggagagccatagtgctcggaaccggtgagatcacacgggaattgcag 180
Dy 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCAG 180
Qy 181 gacgacgggtcccttcttgatcaaacccgctcaatgctgagagattggcgctgcccc 240
Dy 181 GACGACGGGTCTCTTCTTGATCAACCCCGCTCAATGCTCGAGATTTGGCGGTGCCCCC 240
Qy 241 gcgagactgctagccgagtagtgggtgcgaaagccttggtgtagtgcctgtagtagg 300
Dy 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGGCCCTTGTGTTACTGCTGATAGG 300
Qy 301 gtgctgagtagtgcggcgagggtctctgtagacgctgacacgtagacacaaatccataac 360
Dy 301 GTGCTTGCAGTGCCTCCGGGAGGTCTCTGATAGCCGTGCACCATGAGCAGCAATCCTAAAC 360
Qy 351 ctcaagaaaaaacacgtaacacacaccgcccacacagacgctcaagtcccgggcg 420
Dy 351 CTCAGAAAAACCAACAGCTAACACACCCCGCCACACAGGACGCTCAAGTTCCCGGGCG 420
Qy 421 gtgctcagatcgttggtgagtagttacctgttgcgcgcgagggccccaggttggtg 480
Dy 421 GTGCTCAGATCGTTGTTGGAGTTCCTGTTGCGCGCAGGGGCCCGCCAGTTGGGTGTC 480
Qy 481 gcgagactagaaacactccgagcggctcgaaacctcgaggagggagacaaactatccca 540
Dy 481 CGCGATCAGGAAGACTTCCGAGCGGTCCGAACCCCGTGGAGGCGCAGCAGCTATCCCA 540
Qy 541 aggtcgc 600
Dy 541 AGGCTCG 600
Qy 601 gcaacgagggc 660
Dy 601 GCAATGAGGCGATGGGTGGCAGGATGGCTGCTGTCACCCCGCGCGCTCCCGCGCTAGT 660
Qy 651 ggggccccctcgacccccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
Dy 651 GGGGCCCCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 713

RESULT 4

AF207765

LOCUS

AF207766

9379 bp

RNA

VRL

27-APR-2000

DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM

Hepatitis C virus strain MD25 complete genome.

AF207766.1 GI:7650249

Hepatitis C virus.

Hepatitis C virus.

Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

Hepacivirus.

1 (bases 1 to 9379)

Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and

Sato, C.

Characteristics of hepatitis C viral genome associated with disease

progression

unpublished

2 (bases 1 to 9379)

Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and

Sato, C.

Direct Submission

Submitted (23-NOV-1999) Second Department of Internal Medicine,

Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,

Tokyo 113-8519, Japan

Location/Qualifiers

1..9379

/organism="Hepatitis C virus"

/strain="MD25"

/db_xref="taxon:11103"

/note="Isolated from a patient with liver cirrhosis and

hepatocellular carcinoma (HCC-5)"

330..9362

/codon_start=1

/product="polyprotein"

/protein_id="AAF65956.1"

/db_xref="GI:7650250"

/translation="MSTNPKPQRTKRNRRPDVKKPPGGQIVGVYLLPRRGLR

SRWTRKTSRSPQRQPIKARPEGRWAOFGYPWPLYGNEGMAWGLIISPRG

GVNATGPPVPCVRENNSSRCWALTPTLAARNA SVPTKTIIRHVDLLCGAARFSA

VGLDCGSVFLYSQFLTSPRHEVQDCNSIYGHLSGRMANDMMNNSPTAALV

SQLLIPQAMDMVAGAHGVLGLAYFSVGNNAKVLIVMLFAGYDGRVRSGGTA

QRTTVGLARFTTSGPSONIQLVNTNGSHIINRTALNCNDSLOTGFAALFYTHRFNDS

GCPRMASCRIIDKEFGWGPITVAGHSDLPORPCWHYAPQCGIVPAAGCVGVYC

FTSPVVGTSRSGVPTVNGVNESDVLNNTPRPOGNWFGCTWNNSTGFTTCGG

PPCNIIGVNNLTICPTDPRKHPEATVTKCSGSPWLTPRCLVDYPERLWYPCVNE

TILKRVYVGVGEVRLNAACNWRGRCNLEDRSELSPLLSSTYEQWILPCSEFTL

PALSTGLIHQNVDYQLYIGISAVSVFVKWEYVLLFLFLADARVACILMMWLL

IAQAAALEKLVILNAASVAGAHGLSELVFCAAWYIKGKLVPAAYALYGVNPLL

LLALPPLAYANDREMAASCGGVFVGLVLLTSLPHYKVFARLWILWQYLITRAEL

LQWVYVNIIRGGDAIILLTCAAHPELIDISKILLAILGLPLVLOAGITRVPYEV

AOGLIRCMVLRKVAGHYVOMALMKLAALTGTYYVDHLTPLRMAHAGLRDLAVAE

VPFSDMETHVITWCAPIAACVNIISGLPVSARGKEILLGPASHEGRRLAPIT

AYCQTRGLLGCITISLTRDNKNQVEGEVQVSTATQSFATCVNGVCWVIYHGAGSK

TLAQGPBITQMYINVDOLVGLMAPPGARSILPTCTCGSSDLYLIVRHADYIPVRRG

DSRGLSPREVSVLYKSGSGGLPCPSGVVGIIFRAAVCTRGVAKAVDFVPEVSEMTT

TRSPVFTSSPPAVPOTFOVAILHAPHTGVSGKTKVPAAYAAQGVKVLVNPVAAITL

FGFVMSKAGHDVNIIRTVRTITTGAPITVSTYKGLADGCGSGVADYIIMCDECHS

TDSILIGITVLDQAEAGARLVATATPPGSVTVPHPNIERALSNTSEILPEYCK

ALPIEAKGRRHLIFCHSKKCDLAEALSGLGNNAVAYIRGLDVSIVPVSQDVVVA

TDALMTGTDLSDVIDCNTCVTFDFSLDPTFTETFTVPQDAVSRSRRGRTGR

RRGIYRFTYPERSGFSSVLCEDYDAGCAWYELTPTAETSVRLRAYLNTPLGVPCVQ

DHLEFWSVFTGLTHIDAFLSQTKAGDNLPLVAVOATVCAACAPPSWDMQWKC

LIRLPTLGHPTPLLYRLGAVONEVILTHPTTKYIMACMSADLEVTVPWLVGVLA

ALAACTLTSVVLVGRILLSSRPVVDREVLVREFDEMEECASHLPYIQGQOLAE

QFKQKALGLOATASQACAAAPVVEKRWALEAFWAKHWNFIISGIVLAGLISLPGN

PAIASSLMOFTSITSPLOTQSTLILNIIIGWAAQALAPPASAFVQIAGAGAVGSI

GLGVLDILAGYAGVAGALVAFKVMSEMPSTEDVNLPLAIPSLGALVGVVCA

ILLRHVGGEGAVQWNNRLIAFASRNVHSPTVYVPESDAAVRVTOILSSITITQLK

RLHOWINEDCSTPGSGWLRDWMICVTLDFKTLQSKLLPLRPLGVPVYSCQGYK

GVNRGDKLHHTTCGCAQITGVHNGNSHRIVGPRTCSNTHGTFPVAITYTGPCTPSP

APNYSKGLWAAVEYVEVTRGVDFHYVTGMTDNTNVKCPQVPAPEFFTEVDGVRLLR

YAPACKPLREVEITFQVGLNQVLSQLPCEPDVAVLTSMLTDPSHITAEAKRRL

ARGSPPSLSSASQLSAPSKATCTTRHSDPADLIEANLLWROEMGNITRVESEN

KVWILDSPEPLRAEEDEREVSVAEEVLKRSKRPALPIWARPDPYNPPLLESWKDPDY

IPPVVHGCLPPTKAPPIPPRRKRRTVVLTSTVSSALAEALATKTFGSESSAVDSGV
ATAPPDQFDAGDAGDSYESSMPELEGGPDLDSDGWSVTSSEASSEDVWCSSMS
YTWTGALITPCAABESKLPINALSNLLRHNNMVYATTSRSQAQKRVTFDRLQVLD
DHYRGVLKEIKASTVKAALLSVBEACKLPHPHASKRPGYKGRVNLGSKALVNI
RSVMKDLLEDTEPTIDTITMAKNEVEFCVQPKGGRKPARLIVFDLGVCEKMALYD
VVSTLPOAVMGSSYGFQYSPGORVEFLVNAKSKRCPMGFSDYDRCPDSTVENDIRT
TCYIQCDDLAPEARQARLSRTERLYVGGPILNKSQGRHRCRAGSLRVFTAMTRYSCNTL
BSYLKASACRAAKLQDCTMLVCGDDLVVIGESAGTQEDASLRVFTAMTRYSCNTL
DPDQYDELITSCSSNVSAHDASGRVYLLTRDPTTPPLARAAWETARSTPVNSWL
GNITMAYPTLMARMLMTHTFFSILIAQBLKALDCQIYGACYSIEPLDLPQIQRLLH
ELSAFSLHSYSPGEINRVASCLRLGVPLRVRLARSVRKLLSQCGRATCCCKYL
ENMAVTKTKLTPIPAASQOLDNWFVAGYSGGDIYHLSLRARPRFWMCLLLLLSVGV
GIYLLPNR"

BASE COUNT 1853 a 2852 c 2719 g 1955 t

ORIGIN

Query Match 97.1% Score 692; DB 14; Length 9379;

Best Local Similarity 99.3%; Pred. No. 9.4e-167;

Matches 695; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 tggggcgacactcacaccatagatcactccctctgtgaggaactactgtcttcacgcagaa 73

DB 2 TGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAAGTCTTCCACGCAGAA 61

QY 74 agcgtctagccatgagcgttagtagtagtgctgtagcgtccagagcccccctcccggg 133

DB 52 AGCGTCTAGCCATGCGGTTAGTAGTGTGTCGTGTCAGCCTCCAGGACCCCTCCCGG 121

QY 134 agagccatagtggtctgcggaaccggtgagtacaccggaatgccagagcagaccgggtcc 193

DB 122 AGAGCATAGTGTGTCGCGGAACCGGTGAGTAGTACACCGAATTTGCCAGGACACCGGGTCC 181

QY 194 ttcttggaatacaacccgctcaatgctgtagagatttgggctgtcccccgcgagactgtag 253

DB 182 TTCTTGGATCAACCGCTCAATGCTGAGATTGGGCGTCCCGCGGAGACTGCTAG 241

QY 254 ccgagtagtggttgggtgcggaaggccttgtagtactgctgataagggtgcttgagagtg 313

DB 242 CCGAGTAGTGTGGGTGCGGAAGGCTTGTGCTACTGCTGATAGGGTGTTCGCGAGTG 301

QY 314 ccccggaaggtctgtagaccgtgacacatgagcacaaatctcaaacctcaagaaaaac 373

DB 302 CCGCGGAGGTCTCGTAGACCGTGCACCATGAGCAGGATCTTAACCTCAAGAAANAAC 361

QY 374 caaacgtacaacaaacccgcgcacagacgtcaagttcccgggcggtggtgcagactgtag 433

DB 162 CAAAGCTAACACCAACCGCGCCACAGACGTCGAAGTTCCCGGGCGGTGTCAGATCCT 421

QY 434 tggtaggtttaccgtgtgcggaaggccccaggttggtggtgctgcgcgactaggaag 493

DB 422 TGGTGGAGTTTACCTGTTCGCCGCGCAGGGGCCCCAGGTTGGGTGTGCGCGGACTAGGAA 481

QY 494 gacttccagagcgtgcgaacctcgtggaaggcgcaaacctatccccaaagcctcgcccgcc 553

DB 482 GACTTCCGAGCGGTGCGCAACCTCGTGGAGAGCGACACACCTATCCCCAAGGCTCGCCGCC 541

QY 554 cgaggcaggaacctgggtcagcccggtgtagtcttggccctctatggcaacagagggcat 613

DB 542 CGAGGCGAGGACCTGGGCTCAGCCCGGGTATCCTTGGCCCCCTCTATGTCATAGAGGCAT 601

QY 614 ggggtgggcagataggtcctctgcggcgagcccccgagctcccgagcttagtggggcccttcgga 673

DB 602 GGGGTGCGCAGGATGGCTCTCTGCGCCCGTGGCTCCCGCGCTAGTTGGGGCCCCACCGA 661

QY 674 cccccggcgtaggtgcgcgtaatttgggttaaggtcatcgat 713

DB 632 CCCCCGGCGTAGGTGCGGTAAATTTGGGTAAAGTATCATCAT 701

RESULT 5

HCV132996

LOCUS

HCV132996 9418 bp RNA VRL 04-MAY-1999

DEFINITION Hepatitis C virus, complete genome, isolate HCV-AD78.
 ACCESSION AJ132996
 VERSION AJ132996.1 GI:4753718
 KEYWORDS complete genome; polyprotein.
 SOURCE Hepatitis C virus.
 ORGANISM Hepatitis C virus.
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
 REFERENCE 1 (bases 1 to 9418)
 AUTHORS Rispeter, K.
 JOURNAL Thesis (1998) Universitaetsklinikum Essen, Institut fuer Virologie
 REFERENCE 2 (bases 1 to 9418)
 AUTHORS Rispeter, K.
 TITLE Direct Submission
 JOURNAL Submitted (07-MAR-1999) Rispeter K., Universitaetsklinikum Essen, Institut fuer Virologie, Hufelandstrasse 55, 45122 Essen, GERMANY
 FEATURES
 source
 1..9418
 /organism="Hepatitis C virus"
 /viral
 /isolate="HCV-AD78"
 /db_xref="taxon:11103"
 /country="Germany"
 /note="genotype 1b; without entire 3' NCR"
 341..9379
 /codon_start=1
 /product="polyprotein"
 /protein_id="CAB41950.1"
 /db_xref="GI:4753719"
 /translation="MSTNPKPORKTKRNTNRRPDVKEFGGGGIVGGVYLLPRGRPL
 CVARTKTSRSPRRRQDIPKARRPGRGAWAQPYPMPLYNBEGLWAGWLSVRG
 SRVSWPTDRRRNRNLGKVIDLTCGFADLMGYIPLVGAFLGGAARALAHGVRVLED
 GVNATGPGCFSEFLLALLSCLTIPASAYEVNRVSGYVHVNDSCNSIYVETAD
 NMHTGCVPCVRSDNSRCWALPTLPAARNGSVPTAIRRHVDLVLGAAACASMY
 VDCGFSVFLVSLFTSLPRHETVOECNCTYPGHVTHRMAMDMNKNSTALVV
 SOLLRIPOAVDMVAGAHGVLGLAYSMVGNKALLIWMLLFAGVGGGPPRTTIGG
 SQAASGLVSMFSVPSOKIQLINTNGSHHINRTALCNSDSTFGFLAFLVAFN
 SSCPCRMASCRDIPFAQGWPTTAEPSGLDQRPYCHWYAPQCGIPASVSCGPV
 YCTPSPVMTDTRSGVPTIRGNETDVTLLNTRPPQGNWFGTWNNTGFTCTC
 GPPCNGIGLNNLTCTDCTFRKHEATYKCGSGPWLTPRCIVDPYKRLWHYPCV
 NPTIFKRVMGVGEIHLRSAACTNRTGERCNEDRDRSELSPLLSTTEWIIIPCST
 LLTAPSTGLIHLHONIVDQVLYGIGSVLSPAIKWEYILLFLLADARVCACLNM
 LLTAAEALENLVNLAASLAGAHLISFLFFCAAWYIKRLVFGAAYALYSWNL
 LLLLALPRAYAMDREMAASGAVFGLVLLTSLPYKVLAKILWMLYFIRAE
 AHLQVWIPPLNVRGGDAIILLTCAVHEPILFIDITKLLAILGLPLVQLQAIIRAYF
 VRAQGLIRACMLRVAGHYVQMAFMKLAALTGYVYDHLTPLRDWAHAGRLDVA
 ITAYSDTRELIGCIITSLTRDKNOVEGEVQVSTATOSFLATCVNGVQWTVHCG
 VEPVFSQMTKTIITWADTAACDIIILGPLVSARRGKEILLGPADSLGEGWRLAP
 SKTLAGKGPITOMYTNVDOLYGPAPSGARSLTPTCTGSSDLILVSHADVIPRR
 RDSRGLSLPRVSYLKSSGGPILCPSGHVVGIFRAAVCTRGVAKADVFPVESME
 TITRSPVTDNSPPAVPQTFQVHLHAPTSGSKSKTKFAAAYAAQYKVLVLPNSVAA
 TLSFGAYMSKAHGDPSIRTVRTITGAPITYSTYKFLADGGSGGAYDIITCDEC
 HSTDSTSLIGITVLQDAGTAGARLVLTATPPGVTVPHPNIEVALSNTGIEPFY
 GKAPIETIRGGHLLFCFSKKCKDELAQLSGLINAVAYVGLDVSIVPTSGDVVV
 VATDALMTGDTGDSVIDNCVTOTVDESLDPTETIETTVPODAVRSQRRTG
 RGRGRIYRTVTPGERSGDFSDVLCEDDAGCAWELPAETSVRLRYLTPGLPV
 CQHLFEWSFTGLTHIDAHFLSQKQADNFPYLVAQVATVCARAQAPPSWDQGV
 KCLIRPLTHGTPLLYRLGAVQNEVTLTHPVTKYIMACMSADLEAVTSVSDQWQ
 LAAALAYKLTQVATVGRRIILSGKPAIVPDGRVLYREFDEBECAHSLPYIEOGLQ
 AEQFKACGLLQQTAKQAEAAAPVVEKSWRALEAFWAKHMSFISGIOYLAGLITLP
 GNPJIASLMAVTSVPLTTOHTLLFNILGGWVAQALAPPASAAPVAGIAGAANG
 STGLKVLVDILAGYAGVAGALVAEKVMSGEMPTEDLVNLLPALISPLCAVLVGVV
 AALLRVRVGGEGAVGMNRNLATFASRGHNSVPTHYVPESDAARVQILSLITPOL
 LKRLHQMINECSTPCSGSLRDWDWICTVLTDEFTKLQSKLLPLRPGVPFFSCRG
 YKGVWGGDTMOTTCPCGAQITGHVNGSMRIVGPRCTNSNTWHGTFPIINAYTGPCTP
 SPANVFRALRVAAEEYVEVTRVGDVHVVTGMTDNDKPCQVPAPEFTEVDGVL
 RYVAPAKPLLEVEVTFQVGLNQYLVGSOLPCEPEPCVTVLSMLTDPHSHIETAKR
 RLDGSPPLASSASSQSAPSLKATCTTRHDSPPADLLEANLLRWQEMGNITRVES
 ENKVLVDFDPLRAEEREISVPAEILRRKRPFRAMPRIWARPDNFPPLLESWKDP
 DYPVYVHSCGTPPPAPKPPVPPRRRTVLTSTVSSALAEATATPFGSSESSADV
 GTATAPGDSGDPGSDAESCSMPLEGPDPDLDSDGWSVTSVEASDEVCSS
 MSYVTGALITCAAESEKLPINALNSLLRHNMVYATTSSASOROKKVTIFDLQV
 LDDHYWDVLKEMKAKSTYKARLLSVEEACKLTPPHSAKSKFCYGAADVRLNLSKAIN

HINSVWEDLLEDTEPTIDTTIMAKNEVFCVQPEKGGRRKPARLLVFPDLGVVRVCEKMAL
 YDVVSTLPOQVMGSGYFOYSPGORVFLVKAWSSKNPMGFAYDTRCFGSTVTEENDI
 RVGESIYCCDCLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRCRASGVLTSCGN
 TLTYLKASAAACRAAKLODCTMLVCGDVLVIGESAGTOEDAAISLRYFTEAMPYSAP
 PGDPPOPEYDELELITSCSSNVSAHDASGRVYLLTRDPTTPLARAWEARTHTPVNS
 WLGNIIMATPLMARVMVLMTHFFSILLASQLEKALDCQIYGACYSLEPDLDPOLIOR
 LHGLSAPFSHSYSPGEINRVASCLRLKGLVPLPLRVRRHARSVRAKLLSRRGATCGK
 YLFNVAVRTKLLKLTPIPAASRLDSGFVAGYSGGDIYHSLSPARPRFWMCULLLSV
 GVGTYLLPNR"
 BASE COUNT 1861 a 2843 c 2716 g 1998 t
 ORIGIN
 Query-Match 96.9%; Score 691; DB 14; Length 9418;
 Best Local Similarity 98.6%; Pred. No. 1.7e-166;
 Matches 697; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 7 cccctgatggggcgacactccaccatagatcaactccctctgagggaactactctca 66
 DB 6 CCCCTGTTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAACACTGTCTTCA 65
 QY 67 cgcagaagcgtctagccatggcgttagtagatgagtcgtagcagcctccagagaccccc 126
 DB 66 CCGAAGAGCGCTAGCCATGGCGTTTGTATGAGTGTGTCGACGCTCCAGAGCCCCCCC 125
 QY 127 tccgggagagcactagtggtctgcggaaacggtgagtacaccggaattgccaggagcac 186
 DB 126 TCCGGGAGAGCCATAGTGGTCTCGGGAACCGGTGAGTACACCGGAATGCCAGAGCAC 185
 QY 187 cgggtctcttcttggatcaaacccctcaatgcctggagattggcgctgcccccgcgaga 246
 DB 186 CGGTCTCTTTCTTGGATCAACCCCTCAATGCTGTGAGATTTGGCGTGGCCCCCGAGA 245
 QY 247 ctctaaccagtagtttagtcgcaaaagccttctgtagctgctgtagataggtgctt 306
 DB 246 CTGTACCCAGTAGTGTGGTTCGCGAAAGGCGCTTGTGTACTGCTGTATAGGTGCTT 305
 QY 307 gcgagtcggcgaggaggtctctgtagccgtgaccatgagcacaaatcctlaaacctcaaa 366
 DB 306 GCGAGTCCCGCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCTTAACCTCAA 365
 QY 357 gaaaaacaaacgtaaacacacacacgcgcgcacagcgtcaaatctcccggcggtgctt 426
 DB 356 GAAAAACAAACCTTAACACCAACCGCCGCCACAGAGCGTCAAGTTCGCCGGCGGTGCT 425
 QY 427 agatcgttggtggagttactctgttcgcgcagggccccagagttgggtgctgcgcgca 486
 DB 426 AGATCGTGTGGAGTTTACTCTGTTCGCCGCGAGGGCCCCAGGTTGGGTGTGCGCGCA 485
 QY 487 ctagggaagactccgagcgtgctgcaacctgctggaagcgacacactatcccccaagctc 546
 DB 486 CTAGGAAGACTTCGAGCGGTGCGAACCTCGTGAAGCGCACACCTATCCCCAAGCTC 545
 QY 547 gccggccccaggcaggaacctggctcagccccgggtatccttggccctctatggaagc 606
 DB 546 GCCGGCCCCAGGGTAGGGCTGGGCTCAGCCCCGGGTACCTTGGCCCTCTATGGCAACG 605
 QY 607 agggcattggggtggcaggagtgctcctctgcccccgcgctcccgccctagttggggcc 666
 DB 606 AGGCCCTGGGGTGGGACAGGATGCTCTGTCAACCCCGCGGCTCCCGGCCCTAGTTGGGGCC 665
 QY 657 cttcgaccctccgctgagtcgtaatttgggtaaggtcatcgat 713
 DB 656 CCACAGACCCCGCGTAGGTGCGGTAAATTTGGTAAAGGTATCGAT 712
 RESULT 6
 HPC5TRJ4 1880 bp RNA VRL 16-MAY-1998
 LOCUS Hepatitis C virus gene for structural protein, partial cds,
 DEFINITION isolate:HC-J4.
 ACCESSION D00832
 VERSION D00832.1 GI:221513

[illegible]


```

Qy 541 aggtctgcgcggccagggcaggacacgtgggtcagcgggtatctcttgcccccctctatg 600
Db 542 AGGCTGCGCGCGCCGAGGCGAGACCTGGGCTCAGCGCGGGTACCCTTGGCGCCCTCTATG 601

Qy 601 gcaacagggcatggggggcaggatgctcctctgcgcccgcgcgctcccgccctagtt 660
Db 602 GCAACGAGGTCTGGGTGGCGCAGATGCTCTGTCAACCGCAGGCTCCCGGCTAGT 661

Qy 661 ggggcccctcgaccccccgccgctaggtcgctcaatttggttaaggtcatcgat 713
Db 662 GGGGCCCCACGACCGCCGCGTAGTGGCTGCGTAATCTGGGTAAAGTCATCAT 714

RESULT 12
AB049099
LOCUS      AB049099      9504 bp      RNA      VRL      05-DEC-2000
DEFINITION Hepatitis C virus gene for polyprotein, complete cds,
            isolate:HCVT212.
ACCESSION  AB049099
VERSION    AB049099.1 GI:11559464
KEYWORDS   Hepatitis C virus (isolate:HCVT212, specific_host:human) serum cdNA
            to genomic RNA.
ORGANISM   Hepatitis C virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Hepacivirus.
            1 (sites)
            Takahashi,K., Iwata,K., Matsumoto,M., Matsumoto,H., Nakao,K.,
            Hatahara,T., Ohta,Y., Kanai,K., Maruo,H., Baba,K., Hijikata,M. and
            Mishiro,S.
            Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
            with hepatocellular carcinoma: the 'progression score' revisited
            Unpublished
            2. (bases 1 to 9504)
            Mishiro,S.
            Direct Submission
            Submitted (19-SEP-2000) to the DDBJ/EMBL/GenBank databases. Shunji
            Mishiro, Toshiba General Hospital, Department of Medical Sciences;
            6-3-22 Higashi Oh-1, Shinagawa-Ku, Tokyo 140-8522, Japan
            (E-mail:shunji.mishiro@po.toshiba.co.jp, Tel:81-3-3764-8981,
            Fax:81-3-3764-8992)
FEATURES             Location/Qualifiers
     1..9504
        /organism="Hepatitis C virus"
        /isolate="HCVT212"
        /specific_host="human"
        /db_xref="taxon:11103"
        /tissue_type="serum"
        /note="HCV-genotype: 1b"
     1..341
        /note="complete"
     342..9377
        /codon_start=1
        /product="polyprotein"
        /protein_id="BA18812.1"
        /db_xref="GI:11559465"
        /translation="MSTNIPQRKTRNTNRRPDQVKFPGGGIIVGVYLLPRGRPL
GVARTKTSRSOPRRQPIPKARRPEGRWAQPGYWPPLYNEGNGWAGLLSPRG
SRPSWGTDPDRRSRLNGKVIDLTGCFADLMGYIPLVGAPLGAARALAHGVRLVD
GVNATGNDLPGCSGFIPLALLSLCTIPASAYEVNRVNGGYHVTNDCSSNIIYETAD
MIWHTPCGVCPRVRESNYSRGWALTPTLAARNNSVPIMTIRHIDLVLGCTAAFCSAM
VGDLGCSVELYSQLTFTSPRHETVDCNCSITPGHITGHRMADMMNMNNSPTAALVY
SQLRIPOAVVDVAGAHGCVLAGLAYISMVGNWAKVLIWMLFAGVDGRIGLIGTT
SLNTERASFISPGPSOKIOLVNTNGSWHINRTALDPCASLOTGFLAGLLYAKKFNIS
GLERMAASCPIDTFAGGKPIITYTERVNSDQKPYCHVAPRRCGIIPASQVCGPYVC
FTPSPVVGTTDHLGAPTSVRGENGTDLVLLNNTRPPQGNWFGCTWNWSTGFKTCGG
PPCNIIGGGGNTLTQCTDCFRKHPEATYTKCGSPMLTPRCLVDYPIRLWHYPCFTFN
FTLEKVRMYVGVGVEHRLSVACNWTGRCNLEDRDSLSPLLTSTTEWILPCSTFT
LPALSTGLHLHOGIVDQYLYGIGSAFVSFAIKWEYVLLLLFLLDADARVCALMMWL
LIAQAALLENLVNVAAGAHGILSLVFFCAAWYIKGRVPGAAALYAGWVPLL
LLLLPLPRAYAMDREMAACGGAVFGLAFLTSPYKVFARLWLMVLOQYFTRAEA
HLVWVPPNVRGRDAITLLTCAIHPELIFDITLKLLALGLPLMLVLOAGITAVPVFV
RAQLLIRACMLVRKVAGGHVQMALVKLAALTGTYYVDHLTPIQDWAHVALRDLAVV

```

3'UTR

/note="upstream of poly-U stretch"

9416..9504

/note="poly(U) length is indistinct"

poly(U) stretch"

PAGE COUNT 1886 a 2810 c 2683 g 2125 t

ORIGIN

Query Match 96.9%; Score 690.6; DB 14; Length 9504;

Best Local Similarity 98.0%; Pred. No. 2.1e-166;

Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 gccagccccctgatggggggacactccaccatagatcaactccctctgtgaggaactag 60

Db 1 GCCAGCCCCGATTGGGGGGGACACTCCACCATAGATCACTCCCTCTGTGAGGAACACTAG 60

Qy 51 tcttcacagagaagctctagcatggcgttagtatgagtgctgagcagctccagagac 120

Db 61 TCTTTCACGAGAAAGCGCTTAGCCATGGCGTGTACTATGATGTGCTGCGAGCTCCAGGAC 120

Qy 121 ccccccctcccgaggagaccatagtggtctcggaaccggtgaatgacaccgaaatgccag 180

Db 121 CCCCCCTCCCGGAGACCCATAGTGTGCTGCGGAACCGGTGAGTACACCGGAATGGCCAG 180

Qy 131 gaacacccgggtcctttcttgatcaaccccgctcaatgctgagagattggcggtgcccc 240

Db 131 GACGACCGGGTCTTCTTGTGATCAACCCGCTCAATGCTTGAGATTGGCGGTGCCCCC 240

Qy 241 gcgagactgtagccgagtagtggtggtcggaagggccttgctggtactgctcgtatagg 300

Db 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCCCTTGTGTACTGCTCCTGATAGG 300

Qy 301 gtgcttcgagtgccccggagggtctcgtagaccgtgacacatgacacaaatcctaacc 360

Db 301 GTGCTTCGGAGTGCCCCGGGAGGTCTCGTAGACCGCTGCACCATGAGCAGCAATCCTAAC 360

Qy	361	ctcaagaataaaccaaacgtacacacaaacccgcccacaggaagcgtcaagtcccgggcg	420
Db	361	CTCAAGAAAAACCAACCGTAACACCAACCCGCCCCACAGAGAGCTTAAAGTTCCGGGGCG	420
Qy	421	gtgctcagatcgctgggtgagcttaactgtgtgcgcaggggcccagattgggtgtgc	480
Db	421	GTGGTCAGATCGTGGTGGAGCTTACCTGTTGCCGCGCAGGGGCCCTAGGTTGGGTGTGC	480
Qy	481	gcgcgactaggaagactccgagcggtcgcgaacctctggaaggcgacaacctatccccca	540
Db	481	CGCGCACTAGGAAGACTTCCGAGCGGTGCGCAACCTCGTGAAGGCGACAACCTATCCCCA	540
Qy	541	aggctccggcccgagggcgaggacctggctgaagccgggtatcccttggccccctcatg	600
Db	541	AGGCTCGCGCGCCGAGGGCAGGGCCGTGGGCTCAGCCCGGGTACCTTGGCCCGCTACG	600
Qy	601	gcaacgagggcgatgggtgggcaggaatgctctctgcgcccgcgctcccggcctagt	660
Db	601	GCATGAGGGCATGGGTGGCAGAGTGGCTCCTGTACCCCGCGGCTCTGGGCTAGTT	660
Qy	661	ggggcccttcggaccccccgcgctgaggtgcgtgaattgggtgaaggtcatcgat	713
Db	661	GGGGCCCCACGAGACCCCGGGTAGTGGCGCTAATTTGGGTAAAGTCAATCGAT	713
RESULT 13			
-AF333324			
LOCUS	AF333324	9587 bp	mRNA
DEFINITION	Hepatitis C virus type 1b polyprotein mRNA, complete cds.		
ACCESSION	AF333324		15-FEB-2001
VERSION	AF333324.1	GI:12831192	
- KEYWORDS			
SOURCE	Hepatitis C virus type 1b.		
ORGANISM	Hepatitis C virus type 1b		
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1 (bases 1 to 9587)		
AUTHORS	Thomson, M., Nascimbeni, M., Gonzales, S., Murthy, K., Rehmann, B. and Liang, J.		
TITLE	Analyses of viral sequences and virus-specific immune responses during serial passage of an infectious hepatitis C virus serotype 1b clone in chimpanzees		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 9587)		
AUTHORS	Thomson, M. and Liang, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-JAN-2001) LDS, NIDDK/NIH, 10 Center Drive, Bethesda, MD 20892, USA		

```

misc_feature      3031..9404
                  /note="derived from BK strain; similar to sequence
                  deposited in GenBank Accession Number M58335"
3' UTR            9375..9587
misc_feature      9405..9587
                  /note="derived from HCV-la strain H; similar to sequence
                  deposited in GenBank Accession Number AF09070"
BASE COUNT       1921 a 2865 c 2714 g 2087 t
ORIGIN

```

```

Query_Match      96.9%; Score 690.6; DB 14; Length 9587;
Best Local Similarity 98.0%; Pred. No. 2.le-166;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY .1 gccagcccccctgatggggcgagacactccaccatagatcactccccctggaggaaactactg 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db .1 GCCAGCCCCGATGTTGGGGGCACACTCCACCATAGATCACTCCCCCTGTGAGGAAGTACTG 60

QY .61 tcttcacgagaagcgtctagccatggcttagttagtgcgtgcagcctccaggac 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db .61 TCTTTCAGGCAGAAAGCGTCTAGCCATGCGCGTTAGTATGAGTGTGCTGTGCAGCCTCCAGGAC 120

QY 1.21 ccccccctccggggagagcccatagtgtgctgcggaacgggtgagtacacccggaattgccag 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


/protein_id="AAC15722.1"
/db_xref="GI:3098633"
/translation="NSTNPKPQKTKRNTNRRPDQKFPFGGQIGVGVYLLPRGRPL
GYRATKASERSQRRQPIPKARPEGRAMQPGYWPYLYGNEGLGAGWALLSPRG
SRPSKTPDNRNLGKVIDTLCGADLMGYIPLVAGPLGGAARALAGVRLVED
GVNATGNLPGCSFSIFLLALSCITIPASAVEVNSVSIHVINDCNSSIVAEAD
VIMHTPGCPVPCVQEGNSRCWALPTLAARNAVSPTTIRRHVDLLVGTAAFCAMY
VGLGCSIFLSQLETFSPRHETVQDCNSIYPGHSGRHWAMDMNNSPTTALVY
GHTLRPAQVDMVAGHWGLAGLAIYSVMGNKAVLIVALLFGAFGETHTGRVA
SHLTSGFTSLFSSGASQIQLVNTGNSWHINRALKNDLSQTGFALFYAHKFN
CGPERMAKPTIDMAQGWPIIYTKPNSDQRPYCHYAPRCGVPASOVCGVPYC
FPPSPVGTTRDSGPPTYSMENETDMLNTRPQGNWFGCTTMASTGKTKCGG
PCNIGGVNRTLICPTDQFRKHPEATYTKCGSPALTPRCLVDYPRLHWPCPLNF
SIFKVMYGVGVEHLNAACTNWRGRCNLEDRKSELSPLLSTITEWQILPCAFITL
PALSTGLHONIVDVOYLYGVGSFAFVFAIKWEYIILLFLLADARCAAMWMLL
TAQAALAVLLEMAASVAGAGHLSFLVFFCAAMYIKRGLAPGAAYFYGWFWLL
LALLAPPALAYLADREMAASCGGVLGVFLFTLSPIYKVFTRLLIWLQYFTRAEAH
MOVVPLNVRGRDAIILITCAVPELIFDITKLLALLGLPLMWLOAGITRVPYFVR
AOGGLRACMLVRKAGHYVOMVEMKLGALCTGVYVNHITPLRDNAHAGLRDLAVAL
PVFSAMETKVIITWADTAACDIIILGLPVSARRKEIFLGPADSLEGQWRLLAPIT
ATVQQTGRVIGLCITSLDGRKNQVEGEVQVVSSTATQSFATCINGVCMVTVYHAGSK
TLAAGPKGITQMYTNDLDVGNQAPPGARSMTPCSGSSDLYLTRHADYVTPRRRG
DSRGSLSPRPVSYLKGSGGGLLCPSGHVGVFRAAVCTRGVAKAVDFIPVESMETT
MRSPVTNDSTPPAVPTFOVAHLHAPTCSGSKSTKVPAAVAAOGYKVLVLPNSVAATL
CGPAMSKAGHDNPIRTGVRTITTGSTITYSTYCKELADGCGSGAYDILICDECHS
TUSTIILGTVLDAETAGARLVVLATATPPGTVTPHPNIEELGSLNNGEIPPYGK
AIPILAGRRHLIFCHSKKDELAALKTGLGNVAYIRGLDVSVIPPIGVVVVYA
TDALMTGTGDSVDICNTVQTFVDSLDPTFTETTPQDQAVSRQRRTGRC
RSGIVFTVGPGRPMFDSVLCYDAGCAWELTPTAETSVRILRAYLITPGLPVCQ
DLEWESVFTGLTHIDAHFLSOTQAGDNRPYLVAYQATVCAQAAPPSPWDQWKC
LIRKPTLHGTPLLYRLGAVONEVILTHPIKYIMACMSADLEVTSWLVGSGVLA
ALAACLTCSVIVGRILSKPAPVDPREVLYOEFDEMESCASOLPIYEOGMQALAE
PQKALGLQATKOAEAAAPVESKRALETFWAKHMWNLISGIVQIAGLUSTLPGN
PAISALMTASITPLTQNTLNLGWSVAQAAPPASAASFVAGIAGAAGVSI
GLGVLDVTLAGYAGVAGALVAFKVMGEVSTEDLVNLLPAISPGALVGVVCAA
ILRHVGPCEGAVQWNRNLIASFAGNHSVPHYVPEADAAARVQILSLRITOLLK
RUKHINEDSPVPCSGSLWRDVMICVLTDFKTLQSKLLPLRPGVPTFLSCRGYK
GVWGDGMQITCPGCAQIAGHVKNRSMRIVGPTCSNTHGTFPFINATYTPGCPSP
APNSRALNRVAEEVEVTRVDEHYVIGMTIDNKCPCQVPPEFFTEVDGVRLLHR
YAPACKPLREDVTFQVNLQVLSGPCPEPEPVDVTLTSMLTDPDPSHITAKRRL
ARGPPLSSASSASQSLATKCTTHDSDADLIEANLLRWEMGNITRVESN
KVILDSLEPLHAEGDEREISVAEILRKSPFALPIWARPDIYNPLLESWKDPDY
VVPVHGCPPLPTKAPPIPPRRKRTVLTESNVSALAEATKTFGSSGSADVDSY
ATAPDLASDDGDKGDSVSYSSMPLECEPCDPLSDGWSVTSVEEASEDVVCCSMS
YTWCTLIPCAAESEKLPINLSNLLRHNNVYATTSRASLSLQKKKVFTRLOVLD
DHRYDLKMKASTVKAALLSIEACKLTPHSAKSFQVGEKADNRNLSRAVNH
RVSWDELLEDTETPITTIMAKSEVFCQPEKGRKPARLIVFPDLGVTRVSKMALYD
VVSITLPAVMGSSYGFQYSPKQRFVLTNTRSKKCPMGFSYDTRCFDSTVTESDIRV
EESIYOCCLDAPARQAIRSLTERLYIGGLTNSKQNGCYRRCRASGVLTTSCTGNTL
TCYLAKAACRAAKLQDCITMLVNGDVLVVCESAGTOEDAAALRAFTAMTRYSAAPG
DPQPEYDLELITSCSNVSVAHADSGKRYVYLTDRPTPLARAARETARHTPIINSL
GNIIYAPTLMARMLIMTHFFSILLAQEOLKALDCIYAGCISTEPLDPLQIIEHL
GLSATFLHSYSGPEINRVASCLKGLGVPPLTRHARSVRAKLLSQGGRAATCGRYL
FNWAVRTKLTPIPAASQDLDSGWFVAGYSGGDIVYHSLSRARPRWFLCLLLLSVGV
GIYLLPNR"

Db 121 CCCCCCTCCGGGAGAGACCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCAG 180
Qy 181 gacgaccgggtctcttcttgatcaaacccgctcaatgcctggagatttggcgctgcccc 240
Db 181 GAGCACCAGGTCCTCTTTCTTGGATCAACCGCTCAATGCCTGGAGATTTGGCGCTGCCCC 240
Qy 241 gggagactgctagccgagtagtcttgggtgcgaaaggccttggtagctgctgtagg 300
Db 241 GCAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAGAGGCTTGTGTACTGCTGTAGG 300
Qy 301 gtagtgcgagtgccccgggaggtctctgtagaccglgacccatgagcacaatcctaaac 360
Db 301 GTGCTTGGAGTGTCCCGGGGAGGTCGTAGACCGTGCACCATGACGACCAATCCTAAAC 360
Qy 361 ctaaaagaaaaacaaacgtaacacacacgcccgcacacgagcagcgtcaagtccccggcg 420
Db 361 CTCAAGAAAAACAAACGTAACACCAACCGCGCGCACAGGACGTCAGTTCCCGGGCG 420
Qy 421 gtagtcagatcgtagtgaagtttacctgttgcgcgacgagggcccgaggttgggtgc 480
Db 421 GTGGTCAGATCGTTGGTGGAGTTTACCTGTTCGCGCGCAGGGGCGCCAGGTTGGGTGTC 480
Qy 481 ggcgactaggaagacttccgagcggcgcacactcgtggaagcgacacacactatcccca 540
Db 481 GCGCGACTAGGAAGGCTTCCGACGGGTCCGCAACCTCGTGAAGGCGACACCTATCCCA 540
Qy 541 aggtcgccggcccgagggcgagcactgggctcagcccgggtatccttggccccctctatg 600
Db 541 AGGTCGCGCAGCCGAGGGCAGGGCCTGGGCTCAGCCCGGCTACCTTGGCCCCCTATG 600
Qy 601 gcaagagggcatgggggtggggaagatggtcctctgccccggcctccggcctagtt 660
Db 601 GCAATGAGGGCCTGGGGGTGGGAGGATGGCTCTGTACCCCCGCGGCTCCCGGCGCTAGTT 660
Qy 661 ggggccccctcgagccccccggcgttaggtcgcgtaatttgggaaggtcatcgat 713
Db 661 GGGGCCCCAGGACCCCCCGGCGTAGGTCCGCGTAACITGGGTAAAGTCAATCGAT 713

Search completed: March 1, 2002, 15:24:35
Job time: 37284 sec

BASE COUNT 1934 a 2842 c 2698 g 2121 t
-ORIGIN

Query Match 96.9%; Score 690.6; DB 14; Length 9595;
Best Local Similarity 98.0%; Pred. No. 2.1e-166;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 gccagccccctgtagggcgacactccaccatagatcactccccctgtgaggaaactactg 60
Db 1 GCCAGCCCCCTGATGGGGCGACATCCACCATCAATCACTCCCTCTGTAGGAATCTACTG 60

Qy 61 tcttcagcagaagcgttagccatgtagttagtgcgtgcagcctccagagac 120
Db 61 TCTTCAGCAGAAGCGCTAGCCATGGCGTTAGTGTAGTGTGTCGACGCTCCAGGAC 120

Qy 121 cccccctccggagagccatagtcctgcgaacggtgagtcacacggaattgccag 180
Db 121 cccccctccggagagccatagtcctgcgaacggtgagtcacacggaattgccag 180

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 08:33:12 ; Search time 232.06 Seconds
(without alignments)
2634.116 Million cell updates/sec

Title: US-09-763-836-1
Perfect score: 713
Sequence: 1 gccagcccccgtatggggc.....atttgggtaagtcacgat 713

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*

- 1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	713	100.0	713	21	Hepatitis type C v
2	690.6	96.9	9595	20	Infectious hepatitis
3	690.6	96.9	9595	22	Nucleotide sequenc
4	690.6	96.9	9595	22	Infectious hepatitis
5	690.6	96.9	9595	22	Infectious hepatitis
6	689	96.6	780	21	Hepatitis C virus
7	689	96.6	1880	13	Hepatitis C virus
8	689	96.6	2540	13	Hepatitis C virus
9	689	96.6	2540	14	NANBH hepatitis vir
10	689	96.6	2540	15	NANBH genomic fra
11	687.4	96.4	2033	15	Hepatitis C Virus

12	687.4	96.4	2033	16	Hepatitis C virus
13	687.4	96.4	4987	15	Vaccinia virus pro
14	687.4	96.4	4987	16	Vaccinia virus vec
15	685.8	96.2	923	17	Hepatitis C virus
16	684.8	96.0	9472	14	Korean hepatitis C
17	682.6	95.7	7911	13	HCV antigen clone
18	682.4	95.7	9416	13	Non-A, non-B viral
19	682.4	95.7	9416	13	Non-A, non-B viral
20	682.4	95.7	9416	20	Non-A, non-B viral
21	680.8	95.5	3360	17	Hepatitis C genome
22	680.8	95.5	9413	16	Partial HCV non-st
23	679.2	95.3	9413	16	Hepatitis C virus
24	679	95.2	9436	15	Blood transmissibl
25	676	94.8	9413	16	DNA encoding HCV p
26	670.4	94.0	1863	12	Fragment of NANB h
27	668.2	93.7	1554	13	HCV core-envelope
28	667.2	93.6	1734	14	Hepatitis C virus
29	665.6	93.4	1734	14	Hepatitis C virus
30	665.6	93.4	9405	14	Full-length Hepati
31	664	93.1	1734	14	Hepatitis C virus
32	664	93.1	1734	14	Hepatitis C virus
33	663.6	93.1	686	15	Hepatitis C virus
34	663.6	93.1	3461	15	Non-A, non-B hepat
35	663.6	93.1	3461	16	5'UTR/CORE/ENV/NS1
36	661.8	92.8	9401	17	Hepatitis C virus
37	660.8	92.7	1734	14	Hepatitis C virus
38	660.8	92.7	1734	14	Hepatitis C virus
39	660.2	92.6	9416	13	NANBV Hutch c59 is
40	660.2	92.6	9416	19	Hepatitis C virus
41	660	92.6	807	13	PCR-generated frag
42	659.2	92.5	1734	14	Hepatitis C virus
43	659.2	92.5	1734	14	Hepatitis C virus
44	658.8	92.4	3401	15	Non-A, non-B hepat
45	658.8	92.4	3401	16	5'UTR/CORE/ENV/NS1

ALIGNMENTS

RESULT 1
AAA08097 standard; CDNA; 713 BP.
XX
AC AAA08097;
XX
DT 22-JUN-2000 (first entry)
XX
HE Hepatitis type C virus nucleotide sequence SEQ ID NO:1.
KW Hepatitis C virus; HCV; 5' UTR; 5'-nontranslational region; diagnosis;
KW gene expression; infection; IRES; viral; ss.
XX
HE Hepatitis C virus.
XX
Key Location/Qualifiers
FT 5'UTR 1..341
FT CDS 342..713
FT /tag= a
FT /tag= b
FT /note= "no stop codon given"
XX
FN WC200012691.AL.

69 MAR-2000.
08-JUL-1999; 99WO-JP03682.
27-AUG-1998; 98JP-0241367.
(FUSO) FUSO PHARM IND LTD.
Yanada O, Yoshida H, Zhang J;

5b 421 gtggtcagatcgttggtaggttacctgtt

[illegible][illegible]


```
Db 661 ggggccccagaccggcgtaggtgcgttaactgggtaagtcacgat 713
|||||
RESULT 8
AAQ29628
ID AAQ29628 standard; DNA; 2540 BP.
XX
AC AAQ29628;
XX
DT 16-MAR-1993 (first entry)
XX
DE Hepatitis C virus HC-J4 5' region.
XX
KW Non-A non-B hepatitis; NANBH; HCV; detection; diagnosis; screening;
KW PCR; primer; polymerase chain reaction; ss.
XX
OS Hepatitis C virus.
XX
PN EP510952-A.
XX
PD 28-OCT-1992.
XX
PF 23-APR-1992; 92EP-0303625.
XX
PR 26-APR-1991; 91JP-0191376.
XX
PA (IMMO ) IMMUNO JAPAN INC.
XX
PI Nakamura T, Okamoto H;
XX
DR WPI; 1992-359137/44.
XX
PT Detection of non-A, non-B hepatitis virus - using new
PT oligo-nucleotide primers with nucleotide sequences corresp. to
PT part. of the viral RNA
XX
PS Disclosure; Page 18; 54pp; English.
XX
CC This sequence represents the 5' region of hepatitis C virus RNA. The
CC original sample was obtained from human and chimpanzee plasma. RNA
CC was isolated from several samples and homology compared, and the
CC respective sequence of about 1900 - 2500 nucleotides of the 5'
CC terminus and 1100 nucleotides of the 3' terminus determined. The 5'
CC region (given) contains a non-coding region of at least 340
CC nucleotides and a region coding for the structural protein followed
CC by a region coding for the non-structural protein (none actually
CC detailed on the sequence given in the specification). When compared
CC with the sequence of HCV disclosed in EP-388232 this sequence showed
CC homology of 80.5%.
XX
SQ Sequence 2540 BP; 472 A; 775 C; 741 G; 552 T; 0 other;
```

```
Query Match 96.6%; Score 689; DB 13; Length 2540;
Best Local Similarity 97.9%; Pred. No. 1.9e-181;
Matches 698; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 gccagcccccgtatggggcgacactccaccatagatcactccctgtgaggaaactactg 60
Db 1 gccagcccccgtatggggcgacactccaccatagatcactccctgtgaggaaactactg 60
QY 61 tcttcacgcagaaaacgcttagccatgcggttagtatgtagtgcagcctccaggac 120
Db 61 tcttcacgcagaaaacgcttagccatgcggttagtatgtagtgcagcctccaggac 120
QY 121 ccccccctccggagagccatagtgctgcggaacggtagtacacgggaattgccag 180
Db 121 ccccccctccggagagccatagtgctgcggaacggtagtacacgggaattgccag 180
QY 181 gacgacgggtctcttcttgatcacaacccgtccaatgcttgagattggcggtgcccc 240
Db 181 gacgacgggtctcttcttgatcacaacccgtccaatgcttgagattggcggtgcccc 240
```

```
RESULT 9
AAQ43889
ID AAQ43889 standard; cDNA to RNA; 2540 BP.
XX
AC AAQ43889;
XX
DT 21-OCT-1993 (first entry)
XX
DE NANB hepatitis virus polynucleotide N-2540-2.
XX
KW Non-A, non-B; virus; polymerase chain reaction; detection;
KW sensitive; specific; HCV; NANBH; ss.
XX
OS Non-A, non-B hepatitis virus.
XX
FH Key Location/Qualifiers
FT CDS 342..2540
FT /tag= a
FT 5'UTR 1..341
FT /tag= b
FT /note= "from 5' terminal of NANBH virus RNA"
XX
PN JPD5091884-A.
XX
PD 16-APR-1993.
XX
PF 10-APR-1991; 91JP-0196175.
XX
PR 14-JUN-1990; 90JP-0153401.
PR 08-NOV-1990; 90JP-0304405.
XX
PA (NAKA/) NAKAMURA T.
XX
WP1: 1993-199637/25.
DR P-PSDB; AAR38279.
XX
An antigen related to non-A and non-B hepatitis virus - comprises
non-translation region comprising 340 - 341 mols. of nucleotides,
non-translation region comprising 1885 - 2551 mols. of
```

PT: nucleotides including region 1,149 and, etc.

PS Claim 3; Page 19-20; 73pp; Japanese.

The sequence is that of NANB hepatitis virus polynucleotide N-2540-2 which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM. The polypeptide it encodes may be used in a system for detecting NANB hepatitis. This method is highly specific and sensitive, and can detect NANB hepatitis virus which could not be detected by conventional methods.

SQ Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;

Query Match 96.6%; Score 689; DB 14; Length 2540;
Best Local Similarity 97.9%; Pred. No. 1.9e-181;
Matches 698; Conservative 0; Mismatches 15; Indels 0

[illegible]

RESULT 10

RESULT
AAO63753

ID AAQ63753 standard; cDNA to genomic RNA: 2540 BP.

AAQ63753;

XX
1505/55,

•

30 JAN-1995 (first entry)

NA1BVHV genomic fragment #2

Polymerase chain reaction; PCR; primer; amplify; detection: NANBHv; non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss. Synthetic.

JE06125777-A

10-MAY-1994

20-JUN-1991:: 91JP-0247120.

20 JUN-1991: 91JP

(NAKA/) NAKAMURA T.

WPI; 1994-187937/23.

Disclosure: page 24-25; 25pp: Japanese.

The sequences given in AAQ63752-53 represent fragments of the non-A, non-B hepatitis virus (NANBH) genome. These fragments were amplified using the primers given in AAQ63732-51. These primers were used in the detection of NANBH. The primers are based on the 5'-terminal region and the core protein coding region. The method allows highly sensitive detection of NANBH.

Sequence 2540 BP: 470 A: 775 C: 742 G: 553 T: 0 other:

Query Match

Query Match 98.0%; SCORE 689; DB 13;
Best Local Similarity 97.9%; Pred. No. 1.9e-181;

Matches	698	Conservative	0	Mismatches	15	Indels	0	Gaps	0
---------	-----	--------------	---	------------	----	--------	---	------	---

QY	1	gcagccccctgattggtgggagcaactccacatagatcaactccccctgtgaggaaactactg	60
Lb			
Lb	1	gcagagccccgattggtgggagcaactccacatagatcaactccccctgtgaggaaactactg	60
QY	51	tcttcacgacgaagaagcgtctagccatggcgcttagtatgagctgcgtgcagcctccaggac	120
Lb			
Lb	51	tcttcacgacgaagaagcgtctagccatggcgcttagtatgagctgcgtgcagcctccaggac	120
QY	121	ccccctcccggtgagagccaatggtctctcggaacccgttgagtacacgggaattgccag	180
Lb			
Lb	121	ccccctcccggtgagagccaatggtctctcggaacccgttgagtacacgggaattgccag	180
QY	131	gacgacgggtctctcttgattcaaccgcctcaatcctcggagatttggtgcgtgcccc	240
Lb			
Lb	131	gacgacgggtctctcttgattcaaccgcctcaatcctcggagatttggtgcgtgcccc	240
QY	241	gcgagactgctagccgagtagtgttggtgcgcaaaaggccttggtactgcctgatagg	300
Lb			
Lb	241	gcgagactgctagccgagtagtgttggtgcgcaaaaggccttggtactgcctgatagg	300
QY	341	gtctctgcagtagccccgggaggtctcgtagacccgtgcaccatgagcacaaaatccctaaac	360
Lb			
Lb	341	gtctctgcagtagccccgggaggtctcgtagacccgtgcaccatgagcacaaaatccctaaac	360
QY	351	ctcaagaaaaacaaaacgttaacaccaaaccgcgcacagggacgtccaagttcccgggcg	420
Lb			
Lb	351	ctcaagaaaaacaaaacgttaacaccaaaccgcgcacagggacgtccaagttcccgggcg	420
QY	421	gtggtcagatcgttggttgagtttaactgttgccgcaggggccccaggttggtggtgc	480
Lb			
Lb	421	gtggtcagatcgttggttgagtttaactgttgccgcaggggccccaggttggtggtgc	480

Db 1485 gccagcccccctgagtgggggcgacacccaccatagatcactcccctgtgaggaaactactg 1544
QY 61 tcttcacgagaagcgtctagccatggtcggttagttagtgcgtgcagcctccaggac 120
Db 1545 tcttcacgagaagcgtctagccatggtcggttagttagtgcgtgcagcctccaggac 1604
QY 121 ccccccctccggagagaccatagtggtctgcggaacccgtgagtcacacgggaattccag 180
Db 1605 ccccccctccggagagaccatagtggtctgcggaacccgtgagtcacacgggaattccag 1664
QY 181 gacgacgggtccttcttggtgacaccccgctcaatgcctggagatttggtggcgtgcccc 240
Db 1665 gacgacgggtccttcttggtgacaccccgctcaatgcctggagatttggtggcgtgcccc 1724
QY 241 gggagactgctagccagtagtggttggtgcggaagggcctgtgtgtagtgcctgtagg 300
Db 1725 gggagactgctagccagtagtggttggtgcggaagggcctgtgtgtagtgcctgtagg 1784
QY 301 gtccttcgagtgccccggaggtctcgtagaccgtgcaccatgagcacaacatccataac 360
Db 1785 gtccttcgagtgccccggaggtctcgtagaccgtgcaccatgagcacaacatccataac 1844
QY 361 ctcaagaaaaaacacacgtacacacaccccgcccgacagcgctcaagttcccgggcg 420
Db 1845 ccaagaaaaaacacacgtacacacaccccgcccgacagcgctcaagttcccgggcg 1904
QY 421 gtccttcgagtgccccggaggtctcgtagaccgtgcaccatgagcacaacatccataac 480
Db 1905 gtccttcgagtgccccggaggtctcgtagaccgtgcaccatgagcacaacatccataac 1964
QY 481 ggcgactaggaagactcccgagcgtgcgaacctggtggaagggcgacacacatccataac 540
Db 1965 ggcgactaggaagactcccgagcgtgcgaacctggtggaagggcgacacacatccataac 2024
QY 541 aggtctccggcgccggagggcgagcgtggctcagccgggtatccttggtggtggtg 600
Db 2025 aggtctccggcgccggagggcgagcgtggctcagccgggtatccttggtggtggtg 2084
QY 601 gcaacaggggcatgggggtggcaggtggtcctctgctgcgcccggtcccgccgtagtt 660
Db 2085 gcaataggggtgggggtggcaggtggtcctctgctgcgcccggtcccgccgtagtt 2144
QY 661 gggggccttcggaccccgcgtagtcgctgaatttgtaggtcaggtcagat 713
Db 2145 gggggcccgacggccccggcgtagtcgctgaatttgtaggtcaggtcagat 2197

RESULT 14

AAQ86799
ID AAQ86799 standard; DNA; 4987 BP.
XX
AC AAQ86799;
XX
DT
XX
DE
XX
KW Vaccinia virus vector comprising HCV T7N1-19 and firefly luciferase.
KW Vaccinia virus vector; firefly luciferase; hepatitis C virus; T7N1-19;
KW antiviral agent; poliovirus; human rhinovirus;
KW internal ribozyme entry site; non-A non-B; cerebral cardio-hepatitis;
KW foot and mouth disease; ds.
XX
OS Synthetic.
XX
FH Key
FT 1826..4057
FT CDS
XX
XX Location/Qualifiers
XX /*tag= a
XX
XX JP07069899-A.
XX
XX
PD 14-MAR-1995.
XX
XX
PF 02-SEP-1993; 93JP-0241973.

XX 02-SEP-1993; 93JP-0241973.
PR (MITU) MITSUBISHI KASEI CORP.
XX
XX WPI: 1995-144713/19.
DR P-PSDB; AAR72801.
XX
XX Antiviral agent comprising component which disrupts viral gene
PT translation; used for the selective inhibition of e.g. Hepatitis
PT C virus, polio-virus and human rhinovirus
XX
XX
PS Disclosure; Pages 19-22; 23pp; Japanese.
XX
XX AAQ86788 encodes AAR72800 Hepatitis C virus (HCV) T7N1-19, which
CC disrupts viral gene translation, by preventing the binding of
CC the viral mRNA to the internal ribozyme entry site. It was
CC used in the construction of an antiviral agent detecting vector
CC (AAQ86799 which encodes AAR72801), which comprises a vaccinia virus
CC vector, a vector containing T7N1-19 and a firefly luciferase gene.
CC The antiviral agent can be used to treat HCV, poliovirus, cerebral
CC cardio-hepatitis, human rhinovirus and foot and mouth disease viral
CC infections.
XX
XX Sequence 4987 BP; 1516 A; 1050 C; 1081 G; 1322 T; 18 other;
SQ
XX
XX Query Match 96.4%; Score 687.4; DB 16; Length 4987;
XX Best Local Similarity 97.8%; Pred. No. 6.6e-181;
XX Matches 697; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 gccagccccctgagtgggggcgacacccaccatagatcactcccctgtgaggaaactactg 60
Db 1495 gccagccccctgagtgggggcgacacccaccatagatcactcccctgtgaggaaactactg 1544
QY 51 tcttcacgagaagcgtctagccatggtcggttagttagtgcgtgcagcctccaggac 120
Db 1545 tcttcacgagaagcgtctagccatggtcggttagttagtgcgtgcagcctccaggac 1604
QY 121 ccccccctccggagagaccatagtggtctgcggaacccgtgagtcacacgggaattccag 180
Db 1605 ccccccctccggagagaccatagtggtctgcggaacccgtgagtcacacgggaattccag 1664
QY 181 gacgacgggtccttcttggtgacaccccgctcaatgcctggagatttggtggcgtgcccc 240
Db 1665 gacgacgggtccttcttggtgacaccccgctcaatgcctggagatttggtggcgtgcccc 1724
QY 241 ggcgactgctagccagtagtggttggtgcggaagggcctgtgtgtagtgcctgtagg 300
Db 1725 ggcgactgctagccagtagtggttggtgcggaagggcctgtgtgtagtgcctgtagg 1784
QY 301 gtccttcgagtgccccggaggtctcgtagaccgtgcaccatgagcacaacatccataac 360
Db 1785 gtccttcgagtgccccggaggtctcgtagaccgtgcaccatgagcacaacatccataac 1844
QY 361 ctcaagaaaaaacacacgtacacacaccccgcccgacagcgctcaagttcccgggcg 420
Db 1845 ccaagaaaaaacacacgtacacacaccccgcccgacagcgctcaagttcccgggcg 1904
QY 421 gtccttcgagtgccccggaggtctcgtagaccgtgcaccatgagcacaacatccataac 480
Db 1905 gtccttcgagtgccccggaggtctcgtagaccgtgcaccatgagcacaacatccataac 1964
QY 481 ggcgactaggaagactcccgagcgtgcgaacctggtggaagggcgacacacatccataac 540
Db 1965 ggcgactaggaagactcccgagcgtgcgaacctggtggaagggcgacacacatccataac 2024
QY 541 aggtctccggcgccggagggcgagcgtggctcagccgggtatccttggtggtggtg 600
Db 2025 aggtctccggcgccggagggcgagcgtggctcagccgggtatccttggtggtggtg 2084
QY 601 gcaacaggggcatgggggtggcaggtggtcctctgctgcgcccggtcccgccgtagtt 660
Db 2085 gcaataggggtgggggtggcaggtggtcctctgctgcgcccggtcccgccgtagtt 660

Db 2085 gcaatgagggttggtggcgagagtaggtctctgtcaccgccggtcccggtcctagt 2144
AAT28348
ID AAT28348 standard; cDNA; 923 BP.
XX AC AAT28348;
XX DE
XX 22-OCT-1996 (first entry)
XX DE
XX Hepatitis C virus core protein.
XX KW Hepatitis C virus; HCV; hepatitis B virus; HBV; fusion protein;
XX KW immunisation; vaccine; infection; core protein; ds.
XX OS Hepatitis C virus.
XX FH Key Location/Qualifiers
XX CDS 342..917
XX FT /*tag= a
XX FT 5'UTR 1..341
XX FT /*tag= b
XX PN W09610997-A1.
XX PD 18-APR-1996.
XX PF 05-OCT-1995; 95WO-US13552.
XX PR 06-JUN-1995; 95US-04567859.
XX PR 05-OCT-1994; 94US-0318248.
XX PA (APOL-) APOLLON INC.
XX PA (GEHO) GEN HOSPITAL CORP.
XX PI Coney LR, Pachuk CJ, Tokushige K, Wakita T, Wands J;
XX PI Zurawski VR;
XX DR WPI; 1996-209642/21.
XX DR P-PSDB; AAR97018.
XX PT Nucleic acid encoding hepatitis B and C virus fusion proteins - or
XX PT incomplete hepatitis C virus genome, are useful in vaccines for
XX PT prevention or treatment of HBV and HCV infections
XX PS Claim 38; Page 35-36; 53pp; English.
XX CC A new fusion protein comprises a hepatitis B virus (HBV) S gene
XX CC protein (AAT36611) coupled to amino acids 1-69, 1-70 or 1-154 of the
XX CC hepatitis C virus (HCV) core protein (AAT28348).
XX CC The nucleic acid has the coding sequence linked to a CMV promoter,
XX CC RSV enhancer, polyadenylation sequence and opt. the 5'-UTR of HCV.
XX SQ Sequence 923-BP; 170 A; 283 C; 281 G; 189 T; 0 other;

Query Match 96.2%; Score 685.8; DB 17; Length 923;
Best Local Similarity 97.6%; Pred. No. 1.1e-180;
Matches 696; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 gccagcccttgatggggcgacactccacatagatcactccctgtgaggaactactg 60
DB 1 gccagcccttgatggggcgacactccacatagatcactccctgtgaggaactactg 60

QY 61 tcttcacgcagaagcgtctagccatggcgttagtagtagtgcagcctccaggac 120
DB 61 tcttcacgcagaagcgtctagccatggcgttagtagtagtgcagcctccaggac 120

QY 121 ccccccctccggggagagccatagtggtctgcggaacccggtgagtacacccggaattgccag 180
DB 121 ccccccctccggggagagccatagtggtctgcggaacccggtgagtacacccggaattgccag 180

QY 181 gacacccgggtcccttcttcttgatcaaacccgcctcaatccttgagagatttggcggtgccccc 240
DB 181 gacacccgggtcccttcttcttgatcaaacccgcctcaatccttgagagatttggcggtgccccc 240

QY 241 gcgagaccgctagccgagtagtgttggtgcggaagcccttgggtactcctcgatagg 300
DB 241 gcgagaccgctagccgagtagtgttggtgcggaagcccttgggtactcctcgatagg 300

QY 301 gtcttcgagtgccccgggaggtctctgtagaccgtgcacccatgagcacaaatcctaacc 360
DB 301 gtcttcgagtgccccgggaggtctctgtagaccgtgcacccatgagcacaaatcctaacc 360

QY 351 ctcaaaagaaaaaaccaacgtaaacacacacccgcgcacacagagcgtcaagtcccgggcg 420
DB 351 ctcaaaagaaaaaaccaacgtaaacacacacccgcgcacacagagcgtcaagtcccgggcg 420

QY 421 gtggtcagatcgttggtggagtgttacctgttgcgcgcgcaggggccccaggttgggtgtgc 480
DB 421 gtggtcagatcgttggtggagtgttacctgttgcgcgcgcaggggccccaggttgggtgtgc 480

QY 481 gcgcgactaggaagacttcgcgagcgtgcgaacctctgtggaagcgacacacatccccca 540
DB 481 gcgcgactaggaagacttcgcgagcgtgcgaacctctgtggaagcgacacacatccccca 540

QY 541 aggtctgcgcggcccgagggcagacactgggtcgcgcgcgcgggtatccttggccctctatg 600
DB 541 aggtctgcgcggcccgagggcagacactgggtcgcgcgcgcgggtatccttggccctctatg 600

QY 601 gcaacgagggcattgggtggcgaggtgggtcctctgcgcgcgcgggtccccggcctagt 660
DB 601 gcaacgagggcattgggtggcgaggtgggtcctctgcgcgcgcgggtccccggcctagt 660

QY 661 ggggccccttcggacccccgcgtgaggtcgcgtaatttgggttaaggttcacgat 713
DB 661 ggggccccttcggacccccgcgtgaggtcgcgtaatttgggttaaggttcacgat 713

Search completed: March 1, 2002, 15:30:18
Job time: 25026 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 04:58:52 ; Search time 8196.34 Seconds
(without alignments)
934.776 Million cell updates/sec

Title: US-09-763-836-1

Perfect score: 713

Sequence: 1 gcaagccctgtatggggc.....atttgggtgaagtcacgat 713

Scoring table:

IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estfun:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estom:*
- 5: em_estpl:*
- 6: em_estba:*
- 7: em_estro:*
- 8: em_estov:*
- 9: em_htc:*
- 10: gb_est1:*
- 11: gb_est2:*
- 12: gb_htc:*
- 13: gb_gss:*
- 14: em_gss_fun:*
- 15: em_gss_hum:*
- 16: em_gss_inv:*
- 17: em_gss_pln:*
- 18: em_gss_pro:*
- 19: em_gss_rtd:*
- 20: em_gss_vrt:*
- 21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	7.7	1101	13	CNS016UE
2	53.8	7.5	1201	13	CNS016BR
3	51.8	7.3	879	13	CNS02QOG
4	48.4	6.8	925	13	CNS00BNG
5	48.2	6.7	925	13	CNS0091P
6	47.6	6.7	925	13	CNS0091P
7	46.4	6.5	1036	13	CNS010BS
8	44.2	6.2	518	13	CNS016UP
9	44.2	6.2	932	13	CNS0072Q
10	43	6.0	844	13	CNS0052P
11	42.6	6.0	442	10	AF367693
12	42	5.9	932	13	CNS0072Q

C	13	40.4	5.7	384	13	CNS043G6	AL272751 Tetraodon
C	14	40.4	5.7	1203	13	CNS015Y4	AL106054 Drosophil
C	15	40.2	5.6	935	13	CNS006XK	AL106051 Drosophil
C	16	40	5.6	936	11	BG852371	BG852371 1024034A0
C	17	39.6	5.6	487	13	AZ049537	AZ049537 GSSru010
C	18	39.6	5.6	844	13	CNS0052P	AL056652 Drosophil
C	19	39.6	5.6	1009	13	CNS010EW	AL098882 Drosophil
C	20	39.6	5.6	1036	13	CNS010BS	AL098770 Drosophil
C	21	39.6	5.6	1201	13	CNS016BR	AL106545 Drosophil
C	22	39.2	5.5	961	13	CNS01087	AL098641 Drosophil
C	23	39	5.5	1201	13	CNS014BJ	AL103945 Drosophil
C	24	38.8	5.4	977	13	CNS00JX7	AL076850 Drosophil
C	25	38.4	5.4	870	13	CNS006EZ	AL064271 Drosophil
C	26	38	5.3	910	13	CNS006ON	AL065629 Drosophil
C	27	37.8	5.3	1101	13	CNS017RT	AL108419 Drosophil
C	28	37.4	5.2	1101	10	AL513871	AL513871 AL513871
C	29	37.2	5.2	491	13	CNS04QTL	AL303042 Tetraodon
C	30	37.2	5.2	685	11	BG282216	BG282216 602403372
C	31	37.2	5.2	1101	13	CNS017SY	AL108460 Drosophil
C	32	37	5.2	692	13	CNS007WH	AL050923 Drosophil
C	33	37	5.2	723	11	BE967041	BE967041 SP_1035_B
C	34	37	5.2	766	13	AZ197732	AZ197732 SP_1035_B
C	35	36.8	5.2	775	13	AZ184156	AZ184156 SP_1002_B
C	36	36.8	5.2	879	13	CNS02QOG	AL209545 Tetraodon
C	37	36.8	5.2	997	13	CNS006DN	AL065132 Drosophil
C	38	36.8	5.2	1009	13	CNS010EW	AL098882 Drosophil
C	39	36.8	5.2	1101	13	CNS017SY	AL108460 Drosophil
C	40	36.8	5.2	1143	13	AQ897991	AQ897991 HS_3122_B
C	41	36.6	5.1	366	11	R87116	R87116 Yq30h12.s1
C	42	36.6	5.1	1098	13	CNS03F4J	AL241228 Tetraodon
C	43	36.4	5.1	1084	13	CNS05NAK	AL345125 Tetraodon
C	44	36.4	5.1	1100	13	CNS016KD	AL106855 Drosophil
C	45	36.4	5.1	1201	13	CNS01668	AL106346 Drosophil

ALIGNMENTS

RESULT 1

CNS016UE

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CNS016UE 1101 bp DNA GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN16G16 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL107216
AL107216.1 GI:5626210
GSS.
fruit fly.
Plasmid Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
library (DrosBAC) was made by Alain Billaut at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN16G16"


```
BASE COUNT 67 a 317 c 213 g 99 t 183 others
ORIGIN
Query Match 7.3%; Score 51.8; DB 13; Length 879;
Best Local Similarity 41.0%; Pred. No. 0.0052;
Matches 128; Conservative 28; Mismatches 156; Indels 0; Gaps 0;
QY 379 gtaacacaaaccccccacagagcgtcaagttcccgccggtgtgctcagatcgttggtg 438
Db 503 GYBCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 562
QY 439 gaggttacgtgtgcgcgcagggggcccccaggtgtggtgtgcgcgcgactagagaactt 498
Db 563 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 622
QY 499 ccgagcgtgcgaacctcgtggaagcgacaaactatcccccaaggtcgcgcgcgcgagg 558
Db 623 CCCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 682
QY 559 gcaggacctgggtcagccgggtatccttggccctctatgcaacaggggcatggggt 618
Db 683 CGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 742
QY 619 ggcagagtggtcctgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 678
Db 743 GGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 802
QY 679 ggcgtaggtgcg 690
Db 803 CGGGGGGGGGGG 814

RESULT 4
CNS000BNG 1101 bp DNA GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR24B13 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL057398
VERSION AL057398.1 GI:4937864
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
GENSCOPE.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Source
1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR19D16"
```

```
/clone_lib="RPCI-98"
/clone="BACR24B13"
/note="end : TET3"
BASE COUNT 129 a 276 c 132 g 151 t 413 others
ORIGIN
Query Match 6.8%; Score 48.4; DB 13; Length 1101;
Best Local Similarity 25.0%; Pred. No. 0.04;
Matches 62; Conservative 79; Mismatches 107; Indels 0; Gaps 0;
QY 460 ggagcccccaggttggtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 519
Db 701 GGMGMGGMRGKMGKMGKMGKMGKMGKMGKMGKMGKMGKMGKMGKMGKMGKMGKMGK 760
QY 520 gaaggcgacaacctatcccccaaggtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 579
Db 751 CCGKNCCKHCCKCCCKKNCCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCK 820
QY 580 ggtatccttggccctctatggaacagagggcatgggggtggggcagagatgctcctcgc 639
Db 821 CYTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 880
QY 640 cccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 699
Db 881 YCFBKBCCCCKCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 940
QY 700 gtaaggtc 707
Db 941 CKYKGNK 948

RESULT 5
CNS0091P 925 bp DNA GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL053013
VERSION AL053013.1 GI:4934461
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
GENSCOPE.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Source
1. .925
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR19D16"
```

[illegible]

[illegible]

```
BASE COUNT      155 a      202 c      241 g      91 t      243 others
ORIGIN
Query Match
Best Local Similarity 6.2%; Score 44.2; DB 13; Length 932;
Matches 94; Conservative 61; Mismatches 144; Indels 0; Gaps 0;
Qy 332 cgcgccacagacgtcaagtcccgccggtggtgcagatcgttggtgagttacgtgt 451
Dy 922 CGCGCGCCSSSSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 863
Qy 452 gccgcgcaaggcccccaggttggtgtgctgcgcgactagggaagactctccgagcgtcga 511
Dy 862 CSCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 803
Qy 512 acctgtgtaaggcgacacactatcccaaggctcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 571
Dy 802 SCSSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 743
Qy 572 tcagcccggtatcttggccctctatggcaacgagggcgtgggtggcaggtggtgct 631
Dy 742 CGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 683
Qy 632 cctgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 690
Dy 682 SSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 624

RESULT 10
CNS0052P/2P/c      844 bp      DNA      GSS      03-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION      BACRI1P6 of RPCI-98 library from Drosophila melanogaster (fruit
ACCESSION      AL056652
VERSION      AL056652.1 GI:4932342
KEYWORDS      GSS.
SOURCE      fruit fly.
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 844)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osada and
Aaron Mamoser in Piecer de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .844
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACRI1P6"
/note="end : TET3"
/note="TET3"
BASE COUNT      261 a      112 c      92 g      35 t      344 others

Query Match
Best Local Similarity 6.0%; Score 42.6; DB 10; Length 442;
Matches 72; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
Qy 547 gccgcccgcgagggcagcagcctggcctcagcccggtatccttggccctctatggaacg 606
Dy 135 GCCCGTTGATGAGTTGATCTCTCGCTCAGGCATCTCATCTACATAAAAAAATAACG 76
Qy 607 agggcatggggtggcaggtgctcctgtgcgcccgccgctcctggtgggccc 666

BASE COUNT      172 a      121 c      86 g      63 t

Query Match
Best Local Similarity 59.5%; Score 42.6; DB 10; Length 442;
Matches 72; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
Qy 547 gccgcccgcgagggcagcagcctggcctcagcccggtatccttggccctctatggaacg 606
Dy 135 GCCCGTTGATGAGTTGATCTCTCGCTCAGGCATCTCATCTACATAAAAAAATAACG 76
Qy 607 agggcatggggtggcaggtgctcctgtgcgcccgccgctcctggtgggccc 666
```

655 ctagtggggccttcggacccccgcgtaggtcgc 690
796 GGSSSGCGSSSSGSGCGGCGCGSSGSGC 831

Accession	Definition	LOCUS	Size (bp)	Library	Survey	Accession	Version
AL272751.1	Tetraodon nigroviridis genome survey sequence.	AL272751	384	DNA	18-MAY-2000	GI:7994989	

KEYWORDS: CSS; genome survey sequence.
SOURCE: Tetraodon nigroviridis.
ORGANISM: Tetraodon nigroviridis.
CLASSIFICATION: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE: 1 (bases 1 to 384)
AUTHORS: Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Rouzeau, L., Billault, A., Ouetz, F., Saurin, W., Berrot, A., and

WEISSENBACK, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish *Tetraodon nigroviridis*
Unpublished

JOURNAL	2 (bases 1 to 384)
REFERENCE	Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
AUTHORS	Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
	Saurin, W. and Weissenbach, J.
TITLE	Human gene number estimate provided by genome wide analysis using
	Tetradon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 384)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT	This sequence is a single read and was generated as part of a large
	scale clone-end sequencing project of the Tetradon nigroviridis
	genome. For more information, please take a look at
	http://www.genoscope.cns.fr/Tetradon .

```

FEATURES
  source
    1. .384
      /organism="Tetraodon nigroviridis"
      /db_xref="taxon:99883"
      /clone="079J07"
      /clone_lib="G"

```

BASE COORDINATE: 30 3 173 2 78 3 20 4 56 24322

PUC-ori"

/note="Genoscope sequence ID : CUBG079CE04SPI-end ;

Query Match	5.7%	Score 40.4;	DB 13;	Length 384;
Best local Similarity	38.5%	pred.	No. 3.8;	

	Matches	57; Conservative	32; Mismatches	59; Indels	0; Gaps
517	gtgaaggcgacaaacctatcccaaggctccgcggcccgaggcaggacctgggctcagc	576			

214 GCGGAVGSSSSCCCCCCCCSSGGGGGGGRRSSCGGGGGGCGSSCCSSGSSSSCCCGG 155

577 ccggatattcggccccctctatgcaacagggcatggagtgggcagagatgacctgt 636
 154 CCAGGGGGCCCCCCCCCTYCCCCGGSCCGGGGGCGGGGGSGGGSSSSGGSGSGGGG 95
 Db

637 cgcgccgcggctccgcgcctagtgggg 664

Db 94 SC000000SCGSSSSSGGGGYSCSSG 67

RESULTS

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 05:15:42 ; Search time 115.48 Seconds
(without alignments)
1398.327 Million cell updates/sec

Title: US-09-763-836-1

Sequence: 1 gccagccccctgatggggc.....attggtgaagtcacgat 713

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Issue: Patents_NA: *
1: /cgnr2_6/prodata2/1na/5a_COMB.seq.*
2: /cgnr2_6/prodata2/1na/5b_COMB.seq.*
3: /cgnr2_6/prodata2/1na/6a_COMB.seq.*
4: /cgnr2_6/prodata2/1na/6b_COMB.seq.*
5: /cgnr2_6/prodata2/1na/PTORS_COMB.seq.*
6: /cgnr2_6/prodata2/1na/backfiles1.seq.*
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	690.6	96.9	9595	US-09-014-416-4	Sequence 4, Appl1
2	690.6	96.9	9599	US-09-014-416-6	Sequence 6, Appl1
3	689	96.6	780	US-08-474-7008-45	Sequence 45, Appl1
4	685.8	96.2	923	US-08-869-380-1	Sequence 1, Appl1
5	685.8	96.2	923	PCT-US95-13552-14	Sequence 14, Appl1
6	682.4	95.7	1499	US-08-374-977-3	Sequence 3, Appl1
7	682.4	95.7	1499	US-08-384-616-3	Sequence 3, Appl1
8	682.4	95.7	1499	US-08-904-686-3	Sequence 3, Appl1
9	682.4	95.7	1499	US-09-315-850-3	Sequence 3, Appl1
10	682.4	95.7	9416	US-08-324-977-1	Sequence 1, Appl1
11	682.4	95.7	9416	US-08-384-616-1	Sequence 1, Appl1
12	682.4	95.7	9416	US-08-904-686-1	Sequence 1, Appl1
13	682.4	95.7	9416	US-09-315-850-1	Sequence 1, Appl1
14	678.4	95.1	1863	US-08-470-4268-14	Sequence 14, Appl1
15	663.6	93.1	686	US-08-988-3218-37	Sequence 37, Appl1
16	663.6	93.1	686	US-08-397-2208-25	Sequence 25, Appl1
17	661.8	92.8	9401	US-08-432-693-1	Sequence 1, Appl1
18	661.8	92.8	9416	US-08-811-566-19	Sequence 19, Appl1
19	661	92.8	9401	PCT-US91-02225-9	Sequence 9, Appl1
20	659.8	92.5	2116	US-08-101-160-21	Sequence 21, Appl1
21	658.6	92.4	9401	US-07-910-760-9	Sequence 9, Appl1
22	658.6	92.4	9401	US-08-440-519-9	Sequence 9, Appl1
23	658.6	92.4	9599	US-09-014-416-2	Sequence 2, Appl1
24	658.6	92.4	9646	US-08-811-566-1	Sequence 1, Appl1
25	658.6	92.4	12980	US-08-811-566-5	Sequence 5, Appl1
26	643.2	90.2	1863	US-08-470-4268-13	Sequence 13, Appl1
27	639	89.6	9379	US-08-444-818-176	Sequence 176, Appl1

45	553.6	77.9	803	1	US-08-157-233-2	Sequence 2, Appl1
44	555.4	77.9	803	1	US-08-157-233-4	Sequence 4, Appl1
43	557	78.1	803	1	US-08-157-238-4	Sequence 4, Appl1
42	558.6	78.3	803	1	US-08-157-235-6	Sequence 5, Appl1
41	563.4	78.9	803	1	US-08-157-233-5	Sequence 6, Appl1
40	575.8	80.8	9589	1	US-08-157-235-2	Sequence 2, Appl1
39	575.8	80.8	9589	1	US-07-925-695-1	Sequence 1, Appl1
38	577.8	81.0	9511	1	US-07-925-695-7	Sequence 7, Appl1
37	577.8	81.0	9511	1	US-07-925-695-9	Sequence 6, Appl1
36	595.8	83.6	665	3	US-08-444-818-102	Sequence 102, Appl1
35	595.8	83.6	665	3	US-08-444-818-100	Sequence 100, Appl1
34	595.8	83.6	665	3	US-08-444-818-96	Sequence 96, Appl1
33	595.8	83.6	665	3	US-08-444-818-95	Sequence 95, Appl1
32	595.8	83.6	665	3	US-08-444-818-94	Sequence 94, Appl1
31	595.8	83.6	665	3	US-08-444-818-94	Sequence 94, Appl1
30	636.6	89.3	9379	4	US-09-488-87-1	Sequence 1, Appl1
29	636.6	89.3	9185	3	US-08-444-818-133	Sequence 123, Appl1
28	636.6	89.3	9185	3	US-08-444-818-122	Sequence 122, Appl1

ALIGNMENTS

```

RESULT: 1
US-09-014-416-4
Sequence 4, Application US/09014416
Patent No. 6153421
GENERAL INFORMATION:
APPLICANT: Yanagi, Masayuki
APPLICANT: Buhk, Jens
APPLICANT: Emerson, Susanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 20264276
CURRENT APPLICATION NUMBER: US/09/014,416
CURRENT FILING DATE: 1998-01-27
EARLIER APPLICATION NUMBER: US 60/053,062
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ. ID NOS: 65
SOFTWARE: PatentIn Ver. 2.1
SEQ. ID NO. 4
LENGTH: 9595
TYPE: DNA
ORGANISM: Hepatitis C virus
US-09-014-416-4

Query: Match 96.9%; Score 690.6; DB 3; Length 9595;
Best Local Similarity 98.0%; Pred. No. 76-195;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps

Cy 1 gccagcccccctgatgggggcgcacatccacacatagatcactccctgtgaggaactactg 60
Nb 1 gccagcccccctgatgggggcgcacatccacacatgatactccctgtgaggaactactg 60
Cy 1 tcttcacgcagaaagcgtctacgcatggcgcttagatgtagtctctgacggcctccagagc 120
Nb 1 tcttcacgcagaaagcgtctacgcatggcgcttagatgtagtctctgacggcctccagagc 120
Cy 1 ccccccctcccgaggagagcacaatggtctctgcggaaacccggttgatcacaccggattgcag 180
Nb 1 ccccccctcccgaggagagcacaatggtctctgcggaaacccggttgatcacaccggattgcag 180
Cy 1 ccccccctcccgaggagagcacaatggtctctgcggaaacccggttgatcacaccggattgcag 240
Nb 1 ccccccctcccgaggagagcacaatggtctctgcggaaacccggttgatcacaccggattgcag 240
Cy 1 gccagacatgctcttccttccttgatcaaacccgctcaatgctctgagattgggcgtgccccc 300
Nb 1 gccagacatgctcttccttccttgatcaaacccgctcaatgctctgagattgggcgtgccccc 300
Cy 1 gccagacatgctcttccttccttgatcaaacccgctcaatgctctgagattgggcgtgccccc 360
Nb 1 gccagacatgctcttccttccttgatcaaacccgctcaatgctctgagattgggcgtgccccc 360

```

```
Db 301 ggcgttcgagagcccgaggagctcgtatagacgtgcacatgagcagaatccctaac 360
Oy 361 ctcaagaagaacccaacgtaacacccaacgcgcgcacacaggaagcgtacaaatcccgagc 420
Db 361 ctcaagaagaacccaacgtaacacccaacgcgcgcacacaggaagcgtacaaatcccgagc 420
Oy 421 ggggtcagatctgtgtgagtttacctctgtccgcgcaggggcccagaattgggtgtgc 480
Db 421 ggggtcagatctgtgtgagtttacctctgtccgcgcaggggcccagaattgggtgtgc 480
Oy 481 ggcgcactgaggaactctccgagcggtccgaacctctgtgaaggcgacacatcccca 540
Db 481 ggcgcactgaggaactctccgagcggtccgaacctctgtgaaggcgacacatcccca 540
Oy 541 aggttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Db 541 aggttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Oy 601 gcaacgagggcattgggtgaggaatgctctgtccgcgcgcgcgcgcgcgcgcgcgcgcgc 660
Db 601 gcaacgagggcattgggtgaggaatgctctgtccgcgcgcgcgcgcgcgcgcgcgcgcgc 660
Oy 661 ggggccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
Db 661 ggggccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
```

```
RESULT 2
US-09-014-416-6
: Sequence 6, Application US/09014416
: Patent No. 6153421
: GENERAL INFORMATION:
: APPLICANT: Yanagii, Masayuki
: APPLICANT: Buhn, Jens
: APPLICANT: Emerson, Susanne U.
: APPLICANT: Purcell, Robert H.
: TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
: FILE REFERENCE: 20264276
: CURRENT APPLICATION NUMBER: US/09/014.416
: CURRENT FILING DATE: 1998-01-27
: EARLIER APPLICATION NUMBER: US 60/053.062
: EARLIER FILING DATE: 1997-07-18
: NUMBER OF SEQ ID NOS: 65
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 6
: LENGTH: 9599
: TYPE: DNA
: ORGANISM: Hepatitis C virus
US-09-014-416-6
```

```
Query Match 96.98; Score 690.6; DB 3; Length 9599;
Best Local Similarity 98.0%; Pred. NO. 7e-195;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
Oy 1 gccacccccctgagggggcgacactccacatagatcaccctctggaggaactactg 60
Db 1 gccacccccctgagggggcgacactccacatagatcaccctctggaggaactactg 60
Oy 61 tcttaacgacagaagcgtacgacatggtatgtagtctgtcagcctccacagac 120
Db 61 tcttaacgacagaagcgtacgacatggtatgtagtctgtcagcctccacagac 120
Oy 121 cccccctccggagagcgcatagtctgcggaacgcgtgagtaaccggaattgcag 180
Db 121 cccccctccggagagcgcatagtctgcggaacgcgtgagtaaccggaattgcag 180
Oy 181 gacgacgggtctctcttgatcaaccgcgtcaatgcttgagaattgggcggtcccc 240
Db 181 gacgacgggtctctcttgatcaaccgcgtcaatgcttgagaattgggcggtcccc 240
```

```
Oy 241 ggcgactctagccgagtagtctgtgtgcgcgaagcctctgtgtactgctgtatag 300
Db 241 ggcgactctagccgagtagtctgtgtgcgcgaagcctctgtgtactgctgtatag 300
Oy 301 ggcgttcgagagcccgaggagctcgtatagacgtgcacatgagcagaatccctaac 360
Db 301 ggcgttcgagagcccgaggagctcgtatagacgtgcacatgagcagaatccctaac 360
Oy 361 ctcaagaagaacccaacgtaacacccaacgcgcgcacacaggaagcgtacaaatcccgagc 420
Db 361 ctcaagaagaacccaacgtaacacccaacgcgcgcacacaggaagcgtacaaatcccgagc 420
Oy 421 ggggtcagatctgtgtgagtttacctctgtccgcgcaggggcccagaattgggtgtgc 480
Db 421 ggggtcagatctgtgtgagtttacctctgtccgcgcaggggcccagaattgggtgtgc 480
Oy 481 ggcgcactgaggaactctccgagcggtccgaacctctgtgaaggcgacacatcccca 540
Db 481 ggcgcactgaggaactctccgagcggtccgaacctctgtgaaggcgacacatcccca 540
Oy 541 aggttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Db 541 aggttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Oy 601 gcaacgagggcattgggtgaggaatgctctgtccgcgcgcgcgcgcgcgcgcgcgcgcgc 660
Db 601 gcaacgagggcattgggtgaggaatgctctgtccgcgcgcgcgcgcgcgcgcgcgcgcgc 660
Oy 661 ggggccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
Db 661 ggggccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
```

```
RESULT 3
US-08-474-700B-45
: Sequence 45, Application US/08474700B
: Patent No. 6001990
: GENERAL INFORMATION:
: APPLICANT: Wands, Jack
: APPLICANT: Wakita, Takaji
: APPLICANT: Moradpour, Darius
: TITLE OF INVENTION: ANTISENSE INHIBITION OF HEPATITIS C
: TITLE OF INVENTION: VIRUS
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 MB
: COMPUTER: IBM PS/2 Model 502 or 55SX
: OPERATING SYSTEM: MS-DOS (Version 5.1)
: SOFTWARE: WordPerfect (Version 5.1)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474.700B
: FILING DATE: 07-JUN-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/240.382
: FILING DATE: 10 May 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Fraser, Janis K.
: REGISTRATION NUMBER: 34,819
: REFERENCE/DOCKET NUMBER: 00786/279001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 780 base pairs
```


STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333..1499
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..1499
OTHER INFORMATION: /note="sequence = 1 - 1499 of SEQ
OTHER INFORMATION: ID NO: 1"
US-08-324-977-3

Query Match 95.7%; Score 682.4; DB 1; Length 1499;
Best Local Similarity 98.4%; Pred. No. 8,8e-193;
Matches 689; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 14 tgggggagacactccaccatagatacctccctgtgaggaactacgtcttcacggagaa 73
DB 5 tgggggggacactccaccatagatacctccctgtgaggaactacgtcttcacggagaa 64
QY 74 agcctcagcagcgtctgtagtgcgtgcagccctccaggaaccccccctccggg 133
DB 65 AGGCTAGAGCAATGGCGTGTAGTATGATGTCGTGCACACCCCAAGAGACCCCTCCGGG 124
QY 134 agagccatagtggtctgcgggaacgggtgagtacaccggaattccagagacacgggtcc 193
DB 125 AGAGCCATAGTGTGCTGCGGAACCGGTGAGTACACCGGAATTCACAGAGACCGGTCC 184
QY 194 ttcttgagatcaaccggcgtcaatgacctggagatttggcgctgcccccgagagctgtag 253
|||||

DB 185 TTCTTGATCAACCCGCTCATGCTGGAGATTGCGGCTGCCCCCGAGACTGCTAG 214
QY 254 ccgagtagtggtggtgcgaagggccttggtgactcctgtaggtgtcttgagtg 313
DB 245 CCAGTAGTGTGGGTGGCGCAAGGCCCTGTGTGACTGCTGAPAGGCTGCTGGAGTG 304
QY 314 ccccggaaggtctcgttagacgtgacacatgagacaacaaatcctaacaagaaac 373
DB 305 CCGCGGAGCTGCTGTAGACCTGTCACATGACACGAACTCTAAACCTCAAGAAAAAC 364
QY 374 caaacgtaacacacacccgcccacaggaacgtcaagttccggggggtgtgtagtct 433
DB 305 CAACAGTAACACCAACCGCCGCCACAGACGTCAAGTTCCGGGGCTGTGATGCT 424
QY 434 tggtaggttacctgtgtgcgcagggagcccaagttgtgtgtgcgcagactagaa 493
DB 425 TGGTGGAGTTTACTGTGTGCGCGGCAAGGGGCCCAAGTTGGTGTGCGCCGCCAGAA 484
QY 494 gaactccagagcgttcgcaacctcgttgaagggcgaacactatcccaaggtcgcggcc 553
DB 485 GACTTCGAGCGGTGCGCAACCTGTGACAGCGACACACCTATCCCAAGGCTCGCGGCC 544
QY 554 cgaaggaggaacctggtgctcagcccggtatctcttgccctctatgcaagagggat 613
DB 545 CGAGGCGAGACTGGGCTAGCCCGGGTACCTTGAGCTCTGTATGACATAGGACTT 604
QY 614 ggggtggcagagatggtctctgtgcgcccgaggtcccgagctagttggtgaccttgcga 673
DB 605 AGGTGGGCGAGAGAGGCTCTGTACACCCCGGCTCCCGGCTAGTTGGGGGCCCAAGGA 664
QY 674 ccccggtgtagtgcgttaattggttaagttcatgat 713
DB 665 CCGCGGCGTAGGTGCGCTAATTGGGTAGGTCATGAT 704

RESULT 7
US-08-384-616-3
Sequence 3, Application US/08384616
Patent No. 5847101
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKE, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616
CLASSIFICATION: 424
FILING DATE:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990

RESULT 9
US-09-315-850-3
; Sequence 3, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, HIOTO

```

APPLICANT: FUKU, Isao
APPLICANT: Mori, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: Armstrong, Westerman, Hattori, Mcleland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,850
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686
FILING DATE: 01-AUG-1997
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mcleland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333..1499
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1499
OTHER INFORMATION: /note="sequence = 1 - 1499 of SEQ
OTHER INFORMATION: ID NO: 1"
US-09-315-850-3
Query Match 95.7%; Score 682.4; DB 4; Length 1499;
Best Local Similarity 98.4%; Pred. No. 8; 8e-193;

```

Matches 689; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 14 tggggcgacactccacatagatcaactcccttgtaggaactactgtcttcaacgagaa 73
 |||||
 Db 5 tggggcgacactccacatagatcaactcccttgtaggaactactgtcttcaacgagaa 64
 |||||

QY 74 agcgtctagccatggtgtagatagtgctgtgtagcgtccagggccccctcccgag 133
 |||||
 Db 65 agcgtctagccatggtgtagatagtgctgtgtagcgtccagggccccctcccgag 124
 |||||

QY 134 aagacatagtgctgtgtagaaccggttgtagacacgggaattgcccagagacgggtcc 193
 |||||
 Db 125 aagacatagtgctgtgtagaaccggttgtagacacgggaattgcccagagacgggtcc 184
 |||||

QY 194 ttcttgtagtcaaccgctcaatgacctgtagaattggcgctgcccccgagactgtag 253
 |||||
 Db 185 ttcttgtagtcaaccgctcaatgacctgtagaattggcgctgcccccgagactgtag 244
 |||||

QY 254 ccgagtagtggtggtgtagaaggcctgtgtgtactgtcgtatgggtgtgctgtagtg 313
 |||||
 Db 245 ccgagtagtggtggtgtagaaggcctgtgtgtactgtcgtatgggtgtgctgtagtg 304
 |||||

QY 314 ccccggaagctcgtgtagacgctgacacatagacaaatcctlaaacctcaagaanaaac 373
 |||||
 Db 305 ccccggaagctcgtgtagacgctgacacatagacaaatcctlaaacctcaagaanaaac 364
 |||||

QY 374 caacgttaacacacacgcccacacaggaacgacgaatcccggtggtgtagatcgt 433
 |||||
 Db 365 caacgttaacacacacgcccacacaggaacgacgaatcccggtggtgtagatcgt 424
 |||||

QY 434 tggtagatcttaccgtgtgtagcagagggccccaggtgtggtgtgtagcagtagaa 493
 |||||
 Db 425 tggtagatcttaccgtgtgtagcagagggccccaggtgtggtgtgtagcagtagaa 484
 |||||

QY 494 gacttcgagcgtgtgtagacactcgtggaagcgacacacacacacacacacacacac 553
 |||||
 Db 485 gacttcgagcgtgtgtagacactcgtggaagcgacacacacacacacacacacacac 544
 |||||

QY 554 cgaagggcaggaactgtgtagcagcccggtgtagccttccctctatgtagcagagggat 613
 |||||
 Db 545 cgaagggcaggaactgtgtagcagcccggtgtagccttccctctatgtagcagagggat 604
 |||||

QY 614 ggggtgtagcaggaactgtgtagccttcccggtgtagccttccctctatgtagcagag 673
 |||||
 Db 605 ggggtgtagcaggaactgtgtagccttcccggtgtagccttccctctatgtagcagag 664
 |||||

QY 674 ccccggtgtagcgtgtagcgtgtagcgtgtagcgtgtagcgtgtagcgtgtagcgtg 713
 |||||
 Db 665 ccccggtgtagcgtgtagcgtgtagcgtgtagcgtgtagcgtgtagcgtgtagcgtg 704
 |||||

RESULT 10
 US-08-324-977-1
 ; Sequence 1, Application US/08324977
 ; Patent No. 5747339
 GENERAL INFORMATION:
 APPLICANT: OKAYAMA, Hiroto
 APPLICANT: FUKE, Isao
 APPLICANT: MORI, Chisato
 APPLICANT: TAKAMIZAWA, Akahisa
 APPLICANT: YOSHIDA, Iwao
 TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESSSEE: Armstrong, Westernman, Hattori, McLeiland &
 STREET: 1725 K St. N.W. Suite 1000
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20006
 COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/324,977
 FILING DATE: 18-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-167466
 FILING DATE: 25-JUN-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-230921
 FILING DATE: 31-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-305605
 FILING DATE: 09-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/039,706
 FILING DATE: 30-JUL-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/769,996
 FILING DATE: 02-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/635,451
 FILING DATE: 28-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Stevens-Smith, Theresa M.
 REGISTRATION NUMBER: 36,281
 REFERENCE/DOCKET NUMBER: 900703D
 TELEPHONE: (202) 659-2930
 TELEFAX: (202) 887-0357
 TELEX: 440142
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9416 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA from genomic RNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 333..9362
 US-08-324-977-1

Query Match 95.7%; Score 682.4; DB 1; Length 9416;
 Best local Similarity 98.4%; Pred. No. 1,9e-192;
 Matches 689; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 14 tggggcgacactccacatagatcaactcccttgtaggaactactgtcttcaacgagaa 73
 |||||
 Db 5 tggggcgacactccacatagatcaactcccttgtaggaactactgtcttcaacgagaa 64
 |||||

QY 74 agcgtctagccatggtgtagatagtgctgtgtagcgtccagggccccctcccgag 133
 |||||
 Db 65 agcgtctagccatggtgtagatagtgctgtgtagcgtccagggccccctcccgag 124
 |||||

QY 134 aagacatagtgctgtgtagaaccggttgtagacacgggaattgcccagagacgggtcc 193
 |||||
 Db 125 aagacatagtgctgtgtagaaccggttgtagacacgggaattgcccagagacgggtcc 184
 |||||

QY 194 ttcttgtagtcaaccgctcaatgacctgtagaattggcgctgcccccgagactgtag 253
 |||||
 Db 185 ttcttgtagtcaaccgctcaatgacctgtagaattggcgctgcccccgagactgtag 244
 |||||

QY 254 ccgagtagtggtggtgtagaaggcctgtgtgtactgtcgtatgggtgtgctgtagtg 313
 |||||
 Db 245 ccgagtagtggtggtgtagaaggcctgtgtgtactgtcgtatgggtgtgctgtagtg 304
 |||||

QY 314 ccccggaagctcgtgtagacgctgacacatagacaaatcctlaaacctcaagaanaaac 373
 |||||
 Db 305 ccccggaagctcgtgtagacgctgacacatagacaaatcctlaaacctcaagaanaaac 364
 |||||

Sequence 1, Application US/08904686A
Patent No. 5998130
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-0357
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333..9362
US-08-904-686A-1

Query Match 95.7% Score 682.4; DB 2; Length 9416;
Best Local Similarity 98.4%; Pred. No. 1.9e-192;
Matches 689; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 14 tggggcgacactccacatagatcaccctgtgaggaactactgtcttcacgcaagaa 73

|||||
5 TGGGGCGACACTCCACCATAGATCACTCCCTGTGAGCAACTACTGTCTTCACGAGAA 64
QY 74 agcgtctacacatgacgtatagatagtgctgacagctccagaacccccctccggg 133
DB 75 ACCGCTACCACTGGCGTTAGTATGATGCTGTGACGCTCCAGGACCCCCCTCCGGG 124
QY 134 agagccatagtgctctcggaacccggtgagtaaccggaattgcagagacacgggtcc 193
DB 145 AGAGCATAGTGTGCTGCGGAACCGGTGAGTACACGGGAATTGCCAGACGACGGGTCC 184
QY 194 ttcttggatcaaccgctcattgcttgagatgttgggcgtgtcccccgcagactgtctag 253
DB 185 TTTCTTGATCAACCCGCTCAATGCTGAGATTGGCGGTCCCCCGGACAGCTGTAG 244
QY 254 ccgagtagtggtggctcggaaggccttggtactgctgatalagggtgtctggagatg 313
DB 295 CCGAGTAGTGTGGGTGCGGAAGGCTTGTGCTACTGCTATAGGCTGCTTGGAGTG 304
QY 314 ccccggaaggtctgtagacggtgacacatgagcaacaatcctaaacctcaagaagaac 373
DB 305 CCCCGGAGGTCTCGTAGACCGTGCACCATGAGACAGAACTTAACCTCAAGAAAAAC 364
QY 374 caacgtaacaccacacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 433
DB 365 CAACGTAAACACCAACCGCCGACAGACGCTCAAGTTCCGGGCGGTGTGACATGCT 424
QY 414 tggtaggttaccctgtgtcccgcaaggcccggttgggtgctgctgcgcgaactagaa 493
DB 435 TGGTGGAGTTTACCTGTGTCGCCCGCAGGGGCCCGCCAGTTGGTGTGCCGCCACGAGA 484
QY 414 gacttcgagcgtgtcgcaactctgtgaaagcgacaactatcccaaggtctgcggcc 553
DB 435 GACTTCCGAGCGGTGCAACCTCTGGAAGGAGACACTATCCCAAGGCTCGCGGCC 544
QY 554 cgaagcgagacccgtggtcagcccggtatcccttgagcccttatgtgcaagaggaat 613
DB 545 CGAGGGCAGACCTGGCTCAACCCGGGTACCTTGGCTCTCTATGGAAGGAGGCTT 604
QY 614 ggggtgggcaagatlggtctctgtgcccgcggcgtcccggtctgtgggccccttcga 673
DB 615 AGGTGGCAGAGATGGCTCTGTCACCCCGGCGTCCCGGCTATTGGGGCCCGACGGA 664
QY 614 ccccggtgtagtggtggttaattgggtcaaggtcatgat 713
DB 635 CCCCGGAGTGTGCTGCTTAATTGGTAAGTCAATCAT 704
RESULT 13
US-09-315-850-1
Sequence 1, Application US/09315850
Patent No. 6217872
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0

Query Match 95.1%; Score 678.4; DB 2; Length 1863;
 Best Local Similarity 98.4%; Pred. No. 1.5e-191;
 Matches 685; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

QY 18 gggagactcacacataatcctcctctgtagaagactactgtcttaacgcagaagcg 77
   |||
DB 1 GGGGACACTCCACCAATATACCTCCCTGTGAGAACTACTGTTTACGCAAGAACG 60
QY 78 tctagccatggcgttagatgagtgctgtagcctccagagacccctcccgagag 137
   |||
DB 61 TCTAGCCATGGGTTAGTATAGTGTCTGTGAGCCTCCAGAGACCCCTCCGGGAGAG 120
QY 138 ccatagtgctgtagcgaacccggtgtagtaaccggaattgccaagagacccggtcttcc 197
   |||
DB 121 CCATAGTGTGTGCGGAGAACCGGTAGTACACCGGAATTGCCAGAGACCGGCTCTTTC 180
QY 198 ttgagtaaccgcctcaatgctgtagaattggcgctgccccggagaaactgtagccga 257
   |||
DB 181 TTGGATCAACCGGCTCAATGCTGAGATTTGGGCTGCCCGGAGAACTGCTAGCCGA 240
QY 258 gtagtgctgtagcgaagagccctgtgtactgctgtagaggtgctgtagtgcgcgc 317
   |||
DB 241 GTAGTGTGGTGGGAGAAAGGCTGTGTACTGCTGATAGGGTCTTGGAGTGCCTCC 300
QY 318 gggaggtctgtagacgctgtagcgaacaaatcctaaccctcaagaagaaaccaa 377
   |||
DB 301 GGGAGGTCTCTAGAACCTGTGACCATGAGCAAGAACTCTAAACCTCAAGAAACCAAA 360
QY 378 cgtaacacacacccgcgcgcgaagagctcaagttcccgagcggtgtaaatgctgt 437
   |||
DB 361 CGTAAACCAACCGCGCCGCGCAAGAGCTCAAGTTCGCGGCGGTGCTCAATCTTGTGT 420
QY 438 ggaattacactgtgtcgcgcgagagggcccgaggtgtgtgtgtgtgtgtgtgtgt 497
   |||
DB 421 GGAGTTTACTGTGTGCGCGAGAGGGGCCCAAGTTGGGTGTGCGCGGACATGAGAACT 480
QY 498 tccgagcgtgtgcaactcgtgtggaagcgacacactatcccaagagctgcgcgcgcag 557
   |||
DB 481 TCCAGCGCTGTGCAACCTGTGAGTGGAGACACCTATCCCAAGGCTCCGCGAGCCGAG 540
QY 558 ggcagagactggcgtcagccgggtatcctgtgccccctatggaagagagagcag 617
   |||
DB 541 GCGAGCGCTGTGCTACCGCGGTACCTTGGCCCTCTATGGCAATGAGAGGCTTGGGG 600
QY 618 tggcagagatgtctcctgtgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 677
   |||
DB 601 TGGCAGAGATGCTCTGTACACCGCGGCTCCGCGCTAGTGGGGCCCAAGGAGCCG 660
QY 678 cggcgtagtgctgtaatttgggtgaagtcagat 713
   |||
DB 661 CGCGTAGGTGCTGCTACTTGGGTAAAGTCAATCAT 696

```

RESULT 15
 US-08-988-321B-37
 ; Sequence 37, Application US/08988321B
 ; Patent No. 6174868
 ; GENERAL INFORMATION:
 ; APPLICANT: Kevin P. Anderson et al.
 ; TITLE OF INVENTION: Compositions And Methods For Treatment Of Hepatitis C V
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Law Offices of Jane Massey Licata
 ; STREET: 66 East Main Street
 ; City: Marlton
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 08053
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
 ; COMPUTER: IBM COMPATIBLE

OPERATING SYSTEM: WINDOWS 95
 SOFTWARE: WORDPERFECT 6.1 FOR WINDOWS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/988.321B
 FILING DATE: December 10, 1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/650,093
 FILING DATE: May 17, 1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/452,841
 FILING DATE: May 30, 1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/397,220
 FILING DATE: March 9, 1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/945,289
 FILING DATE: September 10, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane Massey Licata
 REGISTRATION NUMBER: 32,257
 REFERENCE/DOCKET NUMBER: ISPH-0245
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (856) 810-1515
 TELEFAX: (856) 810-1454
 INFORMATION FOR SEQ ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 686
 TYPE: nucleic acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 ANTI-SENSE: No
 US-08-988-321B-37

Query Match 93.1%; Score 663.6; DB 4; Length 686;
 Best Local Similarity 80.8%; Pred. No. 2.4e-187;
 Matches 554; Conservative 118; Mismatches 14; Indels 0; Gaps 0;

```

QY 1 gcccaccccgatgagggcgacacatccacatagatcactccctgtgaggaactactg 60
   |||
DB 1 GCCAGCCCCGABUUGGGGGGACACAUCCACCAUAGAUACUCCUUGAGAGACUACUG 60
QY 61 tctcaagcagaagaagcgtctagccatgagcgttagtatagtgctgtcagcctccagac 120
   |||
DB 61 UCUCACCGCAGAAAGCGUCUAGCCAUUGCCGUAGUAGUGUGUCUGUCCAGAC 120
QY 121 ccccccctccggagagagcagatggtgtgtggaacccgtgagtaaccggaaattgacag 180
   |||
DB 121 CCCCCCUCGCGGAGAGGCCAUAAGUGUCUGGAGACCGUGAGUACCGGAUUGGCCAG 180
QY 181 gacgacgggtctcttcttgatcaaccgcctcaatgcttgagagatttgagcgtgcgcc 240
   |||
DB 181 GACGACCGGGUCUCUUGUAGUACACCCGCTCAUUGCUGAGAUUUGGGCGGCCCC 240
QY 241 ggcagactgtcagccagtagtgctgggtgcgcgaagagcctgtgtactgtcagtagg 300
   |||
DB 241 GCGAGACUCGUCAGCCGAGUAGUUGUCUGCCGAAAGGCCUUGUGUAGUCCUGAAGG 300
QY 301 gttcttggaatgcccgggaggtctgtgtagacgtgacacatagagacaaatcctaag 360
   |||
DB 301 GUGGUCUGAGUGGCCCGGGAGGUCUGUAGACCGUGUACCAUAGACGAAUUCUAAAC 360
QY 361 ctcaagaagaaacaaacgtaacacacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
   |||
DB 361 CUCAAAAGAAAACCAACGUAACCAACCAACCGCCGCCACAGAGAGUACUAGUCCGGCG 420
QY 421 gtagtcagtcgtgtgtagttagctgtgtccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 480
   |||
DB 421 GUGGUCAGUACUUGUGUGAGUUAUCCUUGUGCCGCGCAGAGGCCCCACGUGUGGUGUC 480
QY 481 ggcgactaggaaagacttcgagcgcgtgcgaactcgttgaaaggagacaaactatccca 540
   |||

```

```

Db 481 GCCCGAUCAGAGACUUCGAGCGGCCACCCCGUGAAGCGACAGCCUUAUCCCCA 540
Qy 541 aggtctgcccggcccgaaggcaagaccctgggtctagccgggtatctctggccctctatg 600
    ||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 541 AGGCTUGCGCGCGCCGAGGGCAGGGCCUGGGCUCAGCCCGGUAUCCUUGGCCCCUUAUG 600
Qy 601 gcaacgagggcatgggtgtggcagatgtctctgtcgcccgcggtctccggcctagtt 660
    ||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 601 GCADUGAGGGGCAUGGGGUGGGCAGAGUAGGCUCCUACCCCGGCUCCGCGCUAUGU 660
Qy 661 gggggcccttcgagcccccggcgtagg 686
    |||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 661 GGGGCCCAAGGACCCCGCGUAGG 686

```

Search completed: March 1, 2002, 15:26:13
 Job time: 36631 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 15:24:35 ; Search time 13643.2 Seconds

(without alignments)
413.542 Million cell updates/sec

Title: US-09-763-836-7

Perfect score: 342
Sequence: 1 gccagcccccctgatgggggc.....ggctcgttagaccgtgcacc 342

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: 1: gb_ba: 2: gb_hvg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_om: 20: em_or: 21: em_ov: 22: em_pat: 23: em_ph: 24: em_pl: 25: em_ro: 26: em_sts: 27: em_sy: 28: em_un: 29: em_vl: 30: em_hugo_hum: 31: em_hugo_inv: 32: em_hugo_rtd: 33: em_htg_hum: 34: em_htg_inv: 35: em_htg_rtd: 36: em_htg_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	342	100.0	9538	14	AB016785	AB016785 Hepatitis
2	335.6	98.1	9596	14	AF054249	AF054249 Hepatitis
3	328.4	96.0	341	14	HPC5UTR	LA0552 Hepatitis C
4	326.8	95.6	2033	6	E08374	E08374 N-terminal
5	326.8	95.6	2033	6	E08869	E08869 5' region of
6	326.8	95.6	4987	6	E08872	E08872 DNA encodin
7	326.8	95.6	9431	14	HPCPCPO	D45172 Hepatitis C
8	326.8	95.6	9431	14	HPCPP	D30613 Hepatitis C
9	326.8	95.6	9471	6	E06261	E06261 CDNA encodi
10	326.8	95.6	9471	6	E06457	E06457 CDNA encodi
11	325.2	95.1	341	6	AR119856	AR119856 Sequence
12	325.2	95.1	9401	6	AR030378	AR030378 Sequence
13	325.2	95.1	9416	6	AR110845	AR110845 Sequence
14	325.2	95.1	9548	14	D89815	D89815 Hepatitis C
15	325.2	95.1	9595	6	AR119832	AR119832 Sequence
16	325.2	95.1	9595	14	AF054247	AF054247 Hepatitis
17	325.2	95.1	9595	14	AF054248	AF054248 Hepatitis
18	325.2	95.1	9599	6	AR119833	AR119833 Sequence
19	323.6	94.6	341	6	AR119855	AR119855 Sequence
20	323.6	94.6	341	6	AR153763	AR153763 Sequence
21	323.6	94.6	341	14	HPC5NR2	D31602 Human Hepat
22	323.6	94.6	341	14	HPCRNAL	M58406 Hepatitis C
23	323.6	94.6	342	6	AR095001	AR095001 Sequence
24	323.6	94.6	349	14	HCV1A5G	Y10150 Hepatitis C
25	323.6	94.6	350	6	AR028132	AR028132 Sequence
26	323.6	94.6	686	6	AR142355	AR142355 Sequence
27	323.6	94.6	780	6	AR095006	AR095006 Sequence
28	323.6	94.6	923	6	AR153761	AR153761 Sequence
29	323.6	94.6	1880	14	HPC5TRJ4	D00832 Hepatitis C
30	323.6	94.6	2540	6	E04260	E04260 CDNA encodi
31	323.6	94.6	2540	6	E04805	E04805 CDNA to 5'
32	323.6	94.6	2540	6	E07391	E07391 CDNA encodi
33	323.6	94.6	7989	6	AX036255	AX036255 Sequence
34	323.6	94.6	7989	12	SSR242652	AJ242652 Hepatitis
35	323.6	94.6	8001	6	AX036254	AX036254 Sequence
36	323.6	94.6	8001	6	AX036257	AX036257 Sequence
37	323.6	94.6	8001	6	AX036259	AX036259 Sequence
38	323.6	94.6	8001	6	AX036261	AX036261 Sequence
39	323.6	94.6	8001	12	SSR242654	AJ242654 Hepatitis
40	323.6	94.6	8637	6	AX036253	AX036253 Sequence
41	323.6	94.6	8637	12	SSR242651	AJ242651 Hepatitis
42	323.6	94.6	8649	6	AX036256	AX036256 Sequence
43	323.6	94.6	8649	12	SSR242653	AJ242653 Hepatitis
44	323.6	94.6	9401	6	E66593	E66593 Hepatitis C
45	323.6	94.6	9401	6	I71894	I71894 Sequence 9

ALIGNMENTS

RESULT: 1	AB016785	9538 bp	RNA	VRL	28-SEP-1999
LOCUS: 1	AB016785				
DEFINITION: 1	Hepatitis C virus genomic RNA, complete sequence.				
ACCESSION: 1	AB016785				
VERSION: 1	AB016785.1	GI:5821154			
KEYWORDS: 1	polypeptide.				
SOURCE: 1	Hepatitis C virus CDNA to genomic RNA.				
ORGANISM: 1	Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae; Hepacivirus.				
REFERENCE: 1	(sites)				
AUTHORS: 1	Zhang, J., Yamada, O., Ito, T., Akiyama, M., Hashimoto, Y., Yoshida, H.,				
TITLE: 1	Makino, R., Masago, A., Uemura, H. and Ataki, H.				
JOURNAL: 1	A single nucleotide insertion in the 5'-untranslated region of				
MEDLINE: 1	hepatitis C virus leads to enhanced cap-independent translation				
REFERENCE: 1	Virology 261 (2), 263-270 (1999)				
AUTORS: 1	Zhang, J.				

Sept 1, '99

REFERENCE 1 (bases 1 to 2033)

AUTHORS Seki, M., Honda, Y. and Yamada, S.
 TITLE ANTI-SENSE COMPOUND COMPLEMENTARY WITH GENE OF HEPATITIS C VIRUS
 JOURNAL Patent: JP 1994311885-A 1 08-NOV-1994;
 COMMENT MITSUBISHI KASEI CORP
 OS Unknown (hepatitis C virus)
 PN JP 1994311885-A/1
 PD 08-NOV-1994
 PR 09-AUG-1993 JP 1993217095
 PF 25-AUG-1992 JP 92P 248796, 03-MAR-1993 JP 93P 42736 PI
 SEKI MAKOTO, HONDA YOSHIKAZU, YAMADA SUGURU
 PC C12N15/51, A61K39/29, C12N15/11, C12N15/86;
 CC strandedness: Double;
 CC topology: Linear;
 FH key Location/Qualifiers
 FT source 1..2033
 FT /organism='Unclassified'
 FT /note='antislense DNA fragment containing 10 to 34
 FT /note='antislense DNA fragment containing 10 to 34
 FT translation of HCV
 FT gene
 FT complement(250..401)
 FT /note='antislense DNA fragment containing 10 to 34
 FT translation of HCV
 FT nucleotides of this region inhibits FT
 FT translation of HCV
 FT gene
 FT complement(808..859)
 FT /note='antislense DNA fragment containing 10 to 34
 FT nucleotides of this region inhibits FT
 FT translation of HCV
 FT gene
 FT Location/Qualifiers
 FT 1..2033
 FT /organism='Hepatitis C virus'
 FT /db_xref='taxon:11103'
 BASE COUNT 379 a 612 c 595 g 447 t
 ORIGIN

Query Match 95.6%; Score 326.8; DB 6; Length 2033;
 Best Local Similarity 99.1%; Pred. No. 4.5e-82;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 gccagcccccctgattgaggcgacactccacatagatcactccctgtgaggaactactg 60
 Db GCCAGCCCCCTGATGGGGGCGACACTCCACATGATGATCTCCCTGTGAGGAACCTACTG 87
 QY 61 tcttcacgcaagaagcgctctagccatgagcgcttagatagtgctgtagagcctccagcc 120
 Db TCTTCACGCAAGAGCGCTAGCCATGCGCTTAGATAGTGTGCTGACGCTCCAGAC 147
 QY 121 cccccctccggagagacatagtgctgtagcgaacggtagtaacacgggaattgcccag 180
 Db CCCCCCTCCGGAGAGACATAGTGCTGCGGAACCGGTAGTAACACGGGAATTGCCAG 207
 QY 181 gacgacgggtcctctcttgatcaatccgcctcaatgctcgagatttggcgctgcccc 240
 Db GACGACGGGTCCTCTTCTTGATCAA-CCGCTCAATGCCCTGAGATTGGCGCTGCCCC 266
 QY 241 cgcgagactgctagccgagtagtgctggctcgcaaaagccttgtagtgcctgataag 300
 Db CGCGAGACTGCTAGCCCGAGTAGTGTGGCTCGCAAAAGCCTTGCTGACTGCTGATAG 326
 QY 301 ggtccttgagagtgcccgaggaggtctcgtagaccgtgacac 342
 Db GGTCTTGCGAGTGCCCCGGAGGCTCTGTAACCGTGACATC 368
 RESULT 5

E08869
 LOCUS E08869 2033 bp RNA
 DEFINITION 5' region of HCV genome.
 ACCESSION E08869
 VERSION E08869.1 GI:2176973
 KEYWORDS JP 1995069899-A/1.
 SOURCE Hepatitis C virus.
 ORGANISM Hepatitis C virus.
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
 REFERENCE 1 (bases 1 to 2033)
 AUTHORS Seki, M., Honda, Y. and Yamada, S.
 TITLE ANTI-SENSE COMPOUND COMPLEMENTARY WITH GENE OF HEPATITIS C VIRUS
 JOURNAL Patent: JP 1995069899-A 1 14-MAR-1995;
 COMMENT MITSUBISHI CHEM CORP
 OS Unknown (Hepatitis C virus)
 PN JP 1995069899-A/1
 PD 14-MAR-1995
 PF 02-SEP-1993 JP 1993241973
 PI SEKI MAKOTO, HONDA YOSHIKAZU, YAMADA SUGURU
 PC A61K31/70, A61K48/00, C07K7/00, C12N15/09//C07H21/04; CC
 strandedness: Double;
 CC topology: Linear;
 FH key Location/Qualifiers
 FT source 1..2033
 FT /organism='Unclassified'
 FT /clone='77N1-19'
 FT /product='N-terminal region of HCV protein'.
 FEATURES
 SOURCE 1..2033
 /organism='Hepatitis C virus'
 /db_xref='taxon:11103'
 BASE COUNT 379 a 612 c 595 g 447 t
 ORIGIN

Query Match 95.6%; Score 326.8; DB 6; Length 2033;
 Best Local Similarity 99.1%; Pred. No. 4.5e-82;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 gccagcccccctgattgaggcgacactccacatagatcactccctgtgaggaactactg 60
 Db GCCAGCCCCCTGATGGGGGCGACACTCCACATGATGATCTCCCTGTGAGGAACCTACTG 87
 QY 61 tcttcacgcaagaagcgctctagccatgagcgcttagatagtgctgtagagcctccagcc 120
 Db TCTTCACGCAAGAGCGCTAGCCATGCGCTTAGATAGTGTGCTGACGCTCCAGAC 147
 QY 121 cccccctccggagagacatagtgctgtagcgaacggtagtaacacgggaattgcccag 180
 Db CCCCCCTCCGGAGAGACATAGTGCTGCGGAACCGGTAGTAACACGGGAATTGCCAG 207
 QY 181 gacgacgggtcctctcttgatcaatccgcctcaatgctcgagatttggcgctgcccc 240
 Db GACGACGGGTCCTCTTCTTGATCAA-CCGCTCAATGCCCTGAGATTGGCGCTGCCCC 266
 QY 241 cgcgagactgctagccgagtagtgctggctcgcaaaagccttgtagtgcctgataag 300
 Db CGCGAGACTGCTAGCCCGAGTAGTGTGGCTCGCAAAAGCCTTGCTGACTGCTGATAG 326
 QY 301 ggtccttgagagtgcccgaggaggtctcgtagaccgtgacac 342
 Db GGTCTTGCGAGTGCCCCGGAGGCTCTGTAACCGTGACATC 368
 RESULT 6
 E08872
 LOCUS E08872 4987 bp DNA
 DEFINITION DNA encoding a fusion protein of Vaccinia virus protein, HCV protein and Firefly Luciferase.
 ACCESSION E08872

Query Match 95.6%; Score 326.8; DB 14; Length 9431;
 Best Local Similarity 99.1%; Pred. No. 3.5e-82;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

LOCUS E06261 9471 bp RNA PAT 29-SEP-1997
 DEFINITION CDNA encoding genes derived from hepatitis C virus.
 ACCESSION E06261.1 GI:2174448
 VERSION E06261.1
 KEYWORDS JP 1994000085-A/101.
 SOURCE Hepatitis C virus.
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1 (bases 1 to 9471)
 AUTHORS Seki,M., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N.
 GENE OR DNA FRAGMENT DERIVED FROM HEPATITIS C VIRUS, POLYPEPTIDE CODED BY THE SAME AND ITS PRODUCTION
 Patent: JP 1994000085-A 101 11-JAN-1994;
 JOURNAL MITSUBISHI KASEI CORP
 OS (hepatitis C virus)
 PN JP 1994000085-A/101
 PD 11-JAN-1994
 PE 11-JUN-1992 JP 1992194497
 PR 11-JUN-1991 JP 91P 139268, 12-JUL-1991 JP 91P 172794, PR 07-OCT-1991 JP 91P 287008, 16-DEC-1991 JP 91P 332329, PR 20-APR-1992 JP 92P 99957
 PI SEKI MAKOTO, HONDA YOSHIKAZU, TAKAHASHI KAZUNOBU, PI MURAKAMI TOMOKO, TERANISHI YUTAKA, HAYASHI NORIO
 PC C12N15/51,C07K7/06,C07K7/10,C07K13/00,C07K15/12, PC C12N1/21,C12N5/10,
 PC C12N15/11,C12N15/70,C12N15/85,C12P21/02//A61K9/00,A61K9/29, PC (C12N1/21,
 PC C12R1/19),(C12N5/10,C12R1/91),(C12P21/02,C12R1/19),(C12P21/02, PC C12R1/91),
 PC C07K9/00;
 CC strandedness: Double;
 CC topology: Linear;
 CC anti-sense: No;
 CC *source: clone-T7N1-30;
 FH Key Location/Qualifiers
 FT mat_peptide 369..9398
 /product= the peptides reacting specifically

FT and
 FT immunochemically with the serum of hepatitis
 FT type C
 FT patient'.
 FEATURES Location/Qualifiers
 source 1..9471
 BASE COUNT 1892 a 2830 c 2712 g 2037 t
 ORIGIN

Query Match 95.6%; Score 326.8; DB 6; Length 9471;
 Best Local Similarity 99.1%; Pred. No. 3.5e-82;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

LOCUS E06457 9471 bp RNA PAT 29-SEP-1997
 DEFINITION CDNA encoding genes derived from hepatitis C virus.
 ACCESSION E06457.1 GI:2174644
 VERSION E06457.1
 KEYWORDS JP 1994000086-A/101.
 SOURCE Hepatitis C virus.
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1 (bases 1 to 9471)
 AUTHORS Seki,M., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N.
 GENE FOR DNA FRAGMENT DERIVED FROM HEPATITIS C VIRUS, POLYPEPTIDE CODED BY THE SAME AND ITS PRODUCTION
 Patent: JP 1994000086-A 101 11-JAN-1994;
 JOURNAL MITSUBISHI KASEI CORP
 OS (hepatitis C virus)
 PN JP 1994000086-A/101
 PD 11-JAN-1994
 PE 07-OCT-1992 JP 1992293734
 PR 07-OCT-1991 JP 91P 287008, 16-DEC-1991 JP 91P 332329, PR 20-APR-1992 JP 92P 99957
 PI SEKI MAKOTO, HONDA YOSHIKAZU, TAKAHASHI KAZUNOBU, PI MURAKAMI TOMOKO, TERANISHI YUTAKA, HAYASHI NORIO
 PC C12N15/51,C07K7/06,C07K7/10,C07K13/00,C12N5/10, PC C12N15/11,
 PC C12N15/85,C12P21/02//A61K9/29,(C12P21/02,C12R1/91),C07K9/00;
 CC strandedness: Double;
 CC topology: Linear;

CC anti-sense: No;
CC *source: clone=T7N1-30;
FH key Location/Qualifiers
FT mat_peptide 369..9398
FT /product='the peptides reacting specifically
FT and
FT type C
FT immunologically with the serum of hepatitis
FT patient'.
FT Location/Qualifiers
1..9471
/organism='Hepatitis C virus'
/db_xref='taxon:11103'
BASE COUNT 1892 a 2830 c 2712 g 2037 t
ORIGIN

Query Match 95.1%; Score 326.8; DB 6; Length 9471;
Best Local Similarity 99.1%; Pred. No. 3.5e-82;
Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 gccagccccctgattggggcgacactccacatagatccctctgtgaggaactactg 60
DB gccagccccctgattggggcgacactccacatagatccctctgtgaggaactactg 87
QY 61 tcttacgcagaagcgctctagcctatgctgtagatagatgctgtgcaagctcagacc 120
DB tcttacgcagaagcgctctagcctatgctgtagatagatgctgtgcaagctcagacc 147
QY 88 tcttacgcagaagcgctctagcctatgctgtagatagatgctgtgcaagctcagacc 147
DB tcttacgcagaagcgctctagcctatgctgtagatagatgctgtgcaagctcagacc 180
QY 121 ccccccctccggagagacatagtgctcgcgaacccggtgagatgacccggaattgcag 180
DB ccccccctccggagagacatagtgctcgcgaacccggtgagatgacccggaattgcag 207
QY 148 ccccccctccggagagacatagtgctcgcgaacccggtgagatgacccggaattgcag 207
DB ccccccctccggagagacatagtgctcgcgaacccggtgagatgacccggaattgcag 240
QY 181 gagacgcgggtctctctctgataatcccgctcaatgcctgagatgttgagcgtgcacc 240
DB gagacgcgggtctctctctgataatcccgctcaatgcctgagatgttgagcgtgcacc 266
QY 208 gagacgcgggtctctctctgataatcccgctcaatgcctgagatgttgagcgtgcacc 266
DB gagacgcgggtctctctctgataatcccgctcaatgcctgagatgttgagcgtgcacc 300
QY 241 cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 300
DB cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 326
QY 267 cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 326
DB cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 368
QY 301 ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 342
DB ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 368
QY 327 ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 368
DB ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 368

RESULT 11
AR119856 341 bp DNA PAT 16-MAY-2001
LOCUS AR119856 341 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 48 from patent US 6153421.
ACCESSION AR119856
VERSION AR119856.1 GI:14102555
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 341)
AUTHORS Yanagi,M., Bukh,J., Emerson,S.U. and Purcell,R.H.
TITLE Cloned genomes of infectious hepatitis C viruses and uses thereof
JOURNAL Patent: US 6153421-A 48 28-NOV-2000;
FEATURES
source Location/Qualifiers
1..341
/organism='unknown'
BASE COUNT 63 a 103 c 106 g 69 t
ORIGIN

Query Match 95.1%; Score 325.2; DB 6; Length 341;
Best Local Similarity 98.8%; Pred. No. 1.7e-81;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagccccctgattggggcgacactccacatagatccctctgtgaggaactactg 60
DB gccagccccctgattggggcgacactccacatagatccctctgtgaggaactactg 60

DB gccagccccctgattggggcgacactccacatagatccctctgtgaggaactactg 60
QY 61 tcttacgcagaagcgctctagcctatgctgtagatagatgctgtgcaagctcagacc 120
DB tcttacgcagaagcgctctagcctatgctgtagatagatgctgtgcaagctcagacc 120
QY 121 ccccccctccggagagacatagtgctcgcgaacccggtgagatgacccggaattgcag 180
DB ccccccctccggagagacatagtgctcgcgaacccggtgagatgacccggaattgcag 180
QY 181 gagacgcgggtctctctctgataatcccgctcaatgcctgagatgttgagcgtgcacc 240
DB gagacgcgggtctctctctgataatcccgctcaatgcctgagatgttgagcgtgcacc 239
QY 241 cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 300
DB cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 299
QY 240 cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 299
DB cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 299
QY 301 ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 342
DB ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 341
QY 300 ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 341
DB ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 341

RESULT 12
AR030378 9401 bp DNA PAT 29-SEP-1999
LOCUS AR030378 9401 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5861267.
ACCESSION AR030378
VERSION AR030378.1 GI:5943592
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9401)
AUTHORS Su,M.
TITLE Methods, nucleotide sequences and host cells for assaying exogenous
and endogenous protease activity
JOURNAL Patent: US 5861267-A 1 19-JAN-1999;
FEATURES
source Location/Qualifiers
1..9401
/organism='unknown'
BASE COUNT 1868 a 2848 c 2686 g 1999 t
ORIGIN

Query Match 95.1%; Score 325.2; DB 6; Length 9401;
Best Local Similarity 98.8%; Pred. No. 1e-81;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagccccctgattggggcgacactccacatagatccctctgtgaggaactactg 60
DB gccagccccctgattggggcgacactccacatagatccctctgtgaggaactactg 60
QY 61 tcttacgcagaagcgctctagcctatgctgtagatagatgctgtgcaagctcagacc 120
DB tcttacgcagaagcgctctagcctatgctgtagatagatgctgtgcaagctcagacc 120
QY 121 ccccccctccggagagacatagtgctcgcgaacccggtgagatgacccggaattgcag 180
DB ccccccctccggagagacatagtgctcgcgaacccggtgagatgacccggaattgcag 180
QY 181 gagacgcgggtctctctctgataatcccgctcaatgcctgagatgttgagcgtgcacc 240
DB gagacgcgggtctctctctgataatcccgctcaatgcctgagatgttgagcgtgcacc 239
QY 241 cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 300
DB cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 300
QY 240 cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 299
DB cgcgagactgctgacggagtagtggtggtgcgaagagcctgtgtactgctcgtatag 299
QY 301 ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 342
DB ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 341
QY 300 ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 341
DB ggtgctgcgagtgcccgaggaggtcgtgtagaccgtgcacc 341

[illegible][illegible]

QY	1	gcacgccccctgtagggggcgcaacctccacatgatacaaccccccttgggaaactctg	60
Db	1	GCACGCCCCCTGTAGGGGGCGACACTCCACATGTATACCCCCCTTGGAGAACTCTG	60
QY	61	tcttcagcgagaagaagcgtctagaccatggcgttagtagtgtctgacgctccagcc	120
Db	61	TCTTTCAGCGAAGAAAGCCGTCTAGCCCATGGCGTTAATATAGAGTGTGTACACCTCCAGGAC	120
QY	121	ccccctctccggggagagccatagtgctgcggaaccggttgatgacacggaaattgcag	180
Db	121	CCCCCTCCCGGGAGAGGACATAGTGGTGTGGGAACCGGTGTAGTACACCGGAATTGCCAG	180
QY	181	gagacacgggagccttcttcttgatacaatcccgctcaatgacctgagattggcggtgcc	240
Db	181	GACGACCGGGGTCCTTCTTCTTGAGATATA-CCGCGCTAAAGCCCTGGAGATTTGGCGTGC	229
QY	241	cgcgagactgctagccgagtagtgcttggtctcggaagagccttgtgtactgctgtag	300
Db	240	CGCGAGACTGCTTACCCGACAGTGTGTGGGTGCGCAAAAGCGCTTGTGTACTGCCTGATAG	299
QY	301	ggtgcttgcaagtgcccccggaaggtcttcgttagaacgctgaccc	342
Db	300	GGTGCTTTCAGATGCCCGGGAGGATCTCTGTAGACCGTGTCAATC	341

Db 300 GGAGCTTGGAGTGGCCCGGAGAGGTCACGTAAGACCGTGCACC 341

Search completed: March 1, 2002, 15:25:12
Job time: 37321 sec

RESULT	15				
AR119832					
LOCUS					
DEFINITION	AR119832	9595 bp	DNA		
ACCESSION	Sequence	4	from patent	US 6153421.	PAT
VERSION	AR119832				
KEYWORDS	AR119832.1	GI:14102531			
SOURCE					
ORGANISM	Unknown.				
	Unclassified				
REFERENCE	1 (bases 1 to 9595)				
AUTHORS	Yanagi,M., Bukh,J., Emerson,S.U. and Purcell],R.H.				
TITLE	Cloned genomes of infectious hepatitis C viruses and uses thereof				
JOURNAL	Patent: US 6153421-A 4 28-NOV-2000;				
FEATURES	location/Qualifiers				
source	1. 9595				
	/organism="unknown"				
BASE COUNT	1934 a	2842 c	2698 g	2121 t	
ORIGIN					

Query Match	95.1%	Score 325.2	DB 6	Length 9555
Best Local Similarity	98.8%	Pred. No. 1e-01		
Matches 338	Conservative 0	Mismatches 3	Indels 1	Gaps 1
QY	1	gccagccccctgtagtgggagcaactccacatagatcactccctgtgagaactactg	60	
Db	1	gccagccccctgtagtgggagcaactccacatgaaatcacgccctgtgagaaactacg	60	
QY	61	tcttcacgcaaaaagcgtctagacatgcttagatagtgctgtgaagcctcgaagcc	120	
Db	61	tctttacacgcaaaaagcgtcttagccattgacgtttagtatagtctctgacagcctccagac	120	
QY	121	ccccctcccgaggagagcatalgtgtcttggyaaacggtgagtlacacgcgaattgcag	180	
Db	121	ccccctcccgaggagagcattggtgctgcggagacccgtagatcacccgaattgcccag	180	
QY	181	gagacacggagtcctcttcttgatcaatccgcgtcaatgcccggaaatttggtgcgtgcc	240	
Db	181	gagacacggagtcctcttcttgatcaatccgcgtcaatgcccggaaatttggtgcgtgcc	240	
QY	241	cgcgagactgctagccgagtagtgttggtcgcgaaagacctgtgtagtaccctgatag	300	
Db	240	cgcgagactgctagccgagtagtgttggtcgcgaaagacctgtgtagtaccctgatag	299	
QY	301	gtgtccttgagagatgccccggaggtctctcgtagaaccgtgcacc	342	

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

PS Claim 24; Page 64; 94pp; Japanese.

CC The present invention describes an oligonucleotide sequence which
 CC potentiates the expression of a useful gene when incorporated in a
 CC gene expression vector, where the oligonucleotide sequence consists
 CC of the 5'-non-translated region (5'UTR) of a viral genome or its
 CC fragment or modified form. The 5'UTR sequence is useful in screening
 CC potential initiation factors of interaction with the viral IRES or
 CC IRES-dependency, for treatment of diseases caused by cap-dependent
 CC mRNA translation and in the diagnosis of the severity of HCV infection.
 CC The present sequence represents a mutated Hepatitis type C virus (HCV)
 CC 5'UTR nucleotide sequence, which is used in the exemplification of
 CC the present invention.

XX Sequence 342 BP; 62 A; 104 C; 106 G; 70 T; 0 other;

Query Match 100.0%; Score 342; DB 21; Length 342;
 Best Local Similarity 100.0%; Pred. No. 3,3e-92;
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gccagcccccctgtagtggggcgagcactccacatagatcactccctgtgaggaactactg 60
 DB 1 gccagcccccctgtagtggggcgagcactccacatagatcactccctgtgaggaactactg 60
 OY 61 tcttacgcagaagaacgctctagcctatggtatgagtgctgtgcagcctccagagcc 120
 DB 61 tcttacgcagaagaacgctctagcctatggtatgagtgctgtgcagcctccagagcc 120
 OY 121 ccccccctccgggagagacatagtgctgcggaacccgtgtgacacccggaattggccag 180
 DB 121 ccccccctccgggagagacatagtgctgcggaacccgtgtgacacccggaattggccag 180
 OY 181 gacgacccgggtcctctctctgtgataatcccgctcaatgcttgagattggcgctgcccc 240
 DB 181 gacgacccgggtcctctctctgtgataatcccgctcaatgcttgagattggcgctgcccc 240
 OY 241 cgcgagactgctagccgagtagtggtggtcggaagagccttggtgactgctcgatag 300
 DB 241 cgcgagactgctagccgagtagtggtggtcggaagagccttggtgactgctcgatag 300
 OY 301 ggtgctgcgagtgcccgaggagctcgtatagaccgtgcacc 342
 DB 301 ggtgctgcgagtgcccgaggagctcgtatagaccgtgcacc 342

RESULT 2

AAA08097
 ID AAA08097 standard; cDNA: 713 BP.

AC AAA08097;

XX 22-JUN-2000 (first entry)

DE Hepatitis type C virus nucleotide sequence SEQ ID NO:1.

XX Hepatitis C virus; HCV; 5' UTR; 5'-nontranslational region; diagnosis;

KW gene expression; infection; IRES; viral; ss.

OS Hepatitis C virus.

XX Key Location/Qualifiers

FT 5'UTR 1..341

FT CDS /*tag= a

FT /*tag= b

FT /*note= "no stop codon given"

XX W0200012691-A1.

XX 09-MAR-2000.

XX 08-JUL-1999; 99WO-JP03682.

XX 27-AUG-1998; 98JP-0241367.

XX (EJSO) FUSO PHARM IND LTD.

XX Yamada O, Yoshida H, Zhang J;

XX WEI; 2000-237867/20.

XX P-PSDB; AAV82358.

XX Nucleic acid sequence which potentiates gene expression and contains

XX 5'-non-translated region of a viral genome is incorporated in vectors

XX for improving the expression of a useful gene

XX Claim 18; Page 82-84; 94pp; Japanese.

CC The present invention describes an oligonucleotide sequence which
 CC potentiates the expression of a useful gene when incorporated in a
 CC gene expression vector, where the oligonucleotide sequence consists
 CC of the 5'-non-translated region (5'UTR) of a viral genome or its
 CC fragment or modified form. The 5'UTR sequence is useful in screening
 CC potential initiation factors of interaction with the viral IRES or
 CC IRES-dependency, for treatment of diseases caused by cap-dependent
 CC mRNA translation and in the diagnosis of the severity of HCV infection.
 CC The present sequence represents a Hepatitis type C virus (HCV) nucleotide
 CC sequence comprising a 5'UTR, which is used in the exemplification of
 CC the present invention.

XX Sequence 713 BP; 136 A; 218 C; 226 G; 133 T; 0 other;

Query Match 96.0%; Score 328.4; DB 21; Length 713;
 Best Local Similarity 99.4%; Pred. No. 4.3e-88;
 Matches 340; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 gccagcccccctgtagtggggcgagcactccacatagatcactccctgtgaggaactactg 60
 DB 1 gccagcccccctgtagtggggcgagcactccacatagatcactccctgtgaggaactactg 60
 OY 61 tcttacgcagaagaacgctctagcctatggtatgagtgctgtgcagcctccagagcc 120
 DB 61 tcttacgcagaagaacgctctagcctatggtatgagtgctgtgcagcctccagagcc 120
 OY 121 ccccccctccgggagagacatagtgctgcggaacccgtgtgacacccggaattggccag 180
 DB 121 ccccccctccgggagagacatagtgctgcggaacccgtgtgacacccggaattggccag 180
 OY 181 gacgacccgggtcctctctctgtgataatcccgctcaatgcttgagattggcgctgcccc 240
 DB 181 gacgacccgggtcctctctctgtgataatcccgctcaatgcttgagattggcgctgcccc 240
 OY 241 cgcgagactgctagccgagtagtggtggtcggaagagccttggtgactgctcgatag 300
 DB 241 cgcgagactgctagccgagtagtggtggtcggaagagccttggtgactgctcgatag 300
 OY 301 ggtgctgcgagtgcccgaggagctcgtatagaccgtgcacc 342
 DB 301 ggtgctgcgagtgcccgaggagctcgtatagaccgtgcacc 342

RESULT 3

AA064913
 ID AA064913 standard; DNA: 2033 BP.

AC AA064913;

XX 21-DEC-1994 (first entry)

DE Hepatitis C virus genome partial sequence (5'-end).

XX Hepatitis C Virus; Non-A, non-B hepatitis virus; HCV; antisense;

KW therapy; inhibition; viral protein precursor; core protein;

KW envelope protein; non-structural protein; clone 17NI-19; ss.

```
XX Hepatitis C virus.
OS ID A086788 standard; DNA; 2033 BP.
XX AC A086788;
XX FH Key Location/Qualifiers
XX FT 369..2033
XX FT CDS /*tag= a
XX FT /note= "partial open reading frame"
XX CA2104649-A.
XX 26-FEB-1994.
XX 23-AUG-1993; 93CA-2104649.
XX 25-AUG-1992; 93JP-0248796.
XX PR 03-MAR-1993; 93JP-0042736.
XX PA (SEKI/) SEKI M.
XX PI Honda Y, Seki M, Yamada E;
XX WPI; 1994-151836/19.
XX DR P-PSDB; AAR54866.
XX PT Anti-sense oligo:nucleotide(s) complementary to the hepatitis C
XX PT virus genome - are useful as antiviral agents
XX PS Claim 1; Page 59-64; 262pp; English.
XX CC Claimed antisense oligonucleotides have a sequence complementary to
XX CC a base sequence which consists of 10-34 bases, extracted from:
XX CC (1) 93 bases from T at position 107 to A at position 199 of A064913,
XX CC (2) 152 bases from A at position 250 to C at position 401 of
XX CC A064913, or (3) 52 bases from C at position 808 to A at position 859
XX CC of A064913. The HCV sequence A064913 comprises the 5'-end of the
XX CC viral open reading frame, including the core protein coding
XX CC sequence. The new antisense oligonucleotides are useful for
XX CC inhibiting translation of the single polypeptide precursor which is
XX CC subsequently cleaved to produce the HCV viral proteins.
XX SQ Sequence 2033 BP; 379 A; 612 C; 595 G; 447 T; 0 other;

Query Match 95.6%; Score 326.8; DB 15; Length 2033;
Best Local Similarity 99.1%; Pred. No. 1.7e-87;
Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 gccagcccccgtatggtggcgacacccacacatagatcctccctgtgaggaactactg 60
DB 28 gccagcccccgtatggtggcgacacccacacatagatcctccctgtgaggaactactg 87
QY 61 tcttcacgagaagaagcgtctagacatggtctagatgagtgctgcgtgcagccctccagcc 120
DB 88 tcttcacgagaagaagcgtctagacatggtctagatgagtggtgctgcgtgcagccctccagcc 147
QY 121 ccccccctccggagagacatagtggtctgcggaaccggtgagttacacggaaattgcag 180
DB 148 ccccccctccggagagacatagtggtctgcggaaccggtgagttacacggaaattgcag 207
QY 181 gacgacgggtctcttcttgatcaatccgcgtcaatgctgagatttggggtgtccccc 240
DB 208 gacgacgggtctcttcttgatcaatccgcgtcaatgctgagatttggggtgtccccc 266
QY 241 cgcgacagctctagccgagtaagtgtggtctgcggaagcctgtgtactgctatag 300
DB 267 cgcgacagctctagccgagtaagtgtggtctgcggaagcctgtgtactgctatag 326
QY 301 ggtgcttcgagtgcccgaggaggtctcgtagaacgctgacc 342
DB 327 ggtgcttcgagtgcccgaggaggtctcgtagaacgctgacc 368

RESULT 4
```

```
A086788
XX ID A086788 standard; DNA; 2033 BP.
XX AC A086788;
XX FH Key Location/Qualifiers
XX FT 369..2033
XX FT CDS /*tag= a
XX FT /note= "STOP codon absent"
XX JP07069899-A.
XX 14-MAR-1995.
XX 02-SEP-1993; 93JP-0241973.
XX PR 02-SEP-1993; 93JP-0241973.
XX PA (MIRU) MITSUBISHI KASEI CORP.
XX WPI; 1995-144713/19.
XX DR P-PSDB; AAR72800.
XX CC Antiviral agent comprising component which disrupts viral gene
XX CC translation - used for the selective inhibition of e.g. Hepatitis
XX CC C virus, polio-virus and human rhinovirus
XX PS Disclosure: Pages 12-14; 23pp; Japanese.
XX CC A086788 encodes AAR72800 Hepatitis C virus (HCV) T7NI-19, which
XX CC disrupts viral gene translation, by preventing the binding of
XX CC the viral mRNA to the internal ribozyme entry site. It was
XX CC used in the construction of an antiviral agent detecting vector,
XX CC which comprises a vaccinia virus vector, a vector containing
XX CC T7NI-19 and a firefly luciferase gene. The antiviral agent can
XX CC be used to treat HCV, poliovirus, cerebral cardio-hepatitis,
XX CC human rhinovirus and foot and mouth disease viral infections.
XX SQ Sequence 2033 BP; 379 A; 612 C; 595 G; 447 T; 0 other;

Query Match 95.6%; Score 326.8; DB 16; Length 2033;
Best Local Similarity 99.1%; Pred. No. 1.7e-87;
Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 gccagcccccgtatggtggcgacacccacacatagatcctccctgtgaggaactactg 60
DB 28 gccagcccccgtatggtggcgacacccacacatagatcctccctgtgaggaactactg 87
QY 61 tcttcacgagaagaagcgtctagacatggtctagatgagtgctgcgtgcagccctccagcc 120
DB 88 tcttcacgagaagaagcgtctagacatggtctagatgagtgctgcgtgcagccctccagcc 147
QY 121 ccccccctccggagagacatagtggtctgcggaaccggtgagttacacggaaattgcag 180
DB 148 ccccccctccggagagacatagtggtctgcggaaccggtgagttacacggaaattgcag 207
QY 181 gacgacgggtctcttcttgatcaatccgcgtcaatgctgagatttggggtgtccccc 240
DB 208 gacgacgggtctcttcttgatcaatccgcgtcaatgctgagatttggggtgtccccc 266
QY 241 cgcgacagctctagccgagtaagtgtggtctgcggaagcctgtgtactgctatag 300
DB 267 cgcgacagctctagccgagtaagtgtggtctgcggaagcctgtgtactgctatag 300
```

Db 267 cgcgactgctagccgagtagtggtgctgcgaaagccttgctgctgctgctgata 326
 QY 301 ggtgctgcgagtcgcccgaggagctctgttagaccgtgcacc 342
 Db 327 ggtgctgcgagtcgcccgaggagctctgttagaccgtgcacc 368

RESULT 5
 ID AA065322 standard; DNA: 4987 BP.
 AC AA065322;
 DT 21-DEC-1994 (first entry)
 DE Vaccinia virus promoter and HCV core protein-luciferase fusion gene.
 KW Hepatitis C Virus; HCV; control; Non-A, non-B hepatitis virus;
 KW antisense; therapy; inhibition; viral protein precursor;
 KW recombinant vaccinia virus; HCV core protein gene; firefly;
 KW luciferase reporter gene; fusion construct; ds.
 OS Chimeric Vaccinia virus.
 OS Chimeric Hepatitis C virus.
 OS Chimeric Photinus pyralis.
 XX
 FH Key Location/Qualifiers
 FT CDS 1826..4057
 FT /*tag= a
 FT /note= "encodes HCV core protein-luciferase
 fusion protein"
 XX
 PN CA2104649-A.
 PD 26-FEB-1994.
 PD XX
 PD 23-AUG-1993; 93CA-2104649.
 PF XX
 PF 25-AUG-1992; 92JP-0248796.
 PR 03-MAR-1993; 93JP-0042736.
 XX
 PA (SEKI/) SEKI M.
 XX
 PI Honda Y, Seki M, Yamada E;
 PI XX
 DR WPI: 1994-151836/19.
 DR P-PSDB: AAR34867.
 XX
 XX
 PT Anti-sense oligo:nucleotide(s) complementary to the hepatitis C
 PT virus genome - are useful as antiviral agents
 XX
 PS Example 5; Page 243-251; 262pp; English.
 XX
 CC A recombinant vaccinia virus which codes for a HCV core protein
 CC fused to the firefly luciferase enzyme was constructed from PCR
 CC amplified fragments. The construct was useful for assaying the
 CC inhibitory activity of various antisense oligonucleotides on HCV
 CC gene translation.
 CC
 XX
 SQ Sequence 4987 BP; 1520 A; 1052 C; 1083 G; 1332 T; 0 other;

Query Match 95.6%; Score 326.8; DB 15; Length 4987;
 Best Local Similarity 99.1%; Pred. No. 2.1e-87;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ggcggcccccgcgagtgaggagcactccacatagatcacccttgaggaaactactg 60
 Db 1485 ggcggcccccgcgagtgaggagcactccacatagatcacccttgaggaaactactg 1544

QY 61 tcttcacgcagaaagcgtcttagcattgagtgatgagtgctgcagcctccagcc 120
 Db 1545 tcttcacgcagaaagcgtcttagcattgagtgatgagtgctgcagcctccagcc 1604

QY 121 ccccccctccgggagagcattgctgctgcgaaacgggtgatacccggaattgcag 180
 Db 1605 ccccccctccgggagagcattgctgctgcgaaacgggtgatacccggaattgcag 1664

QY 181 gacgacgggtctcttcttgatcaatccgcgtcaatgctggaattggcgctgcc 240
 Db 1665 gacgacgggtctcttcttgatcaatccgcgtcaatgctggaattggcgctgcc 1723

QY 241 cgcgagactgctagccgagtagtggtgctgcgaaagccttgctgactgctgata 300
 Db 1724 cgcgagactgctagccgagtagtggtgctgcgaaagccttgctgactgctgata 1783

QY 301 ggtgctgcgagtcgcccgaggagctctgttagaccgtgcacc 342
 Db 1794 ggtgctgcgagtcgcccgaggagctctgttagaccgtgcacc 1825

RESULT 6
 ID AA086799 standard; DNA: 4987 BP.
 AC AA086799;
 DT 04-DEC-1995 (first entry)
 DE Vaccinia virus vector comprising HCV T7N1-19 and firefly luciferase.
 KW Vaccinia virus vector; firefly luciferase; hepatitis C virus; T7N1-19;
 KW antiviral agent; poliovirus; human rhinovirus;
 KW internal ribozyme entry site; non-A non-B; cerebral cardio-hepatitis;
 KW foot and mouth disease; ds.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1826..4057
 FT /*tag= a
 XX
 PN JP07069899-A.
 PD 14-MAR-1995.
 PD XX
 PD 02-SEP-1993; 93JP-0241973.
 PF XX
 PF 02-SEP-1993; 93JP-0241973.
 PR XX
 PR 02-SEP-1993; 93JP-0241973.
 XX
 PA (MITU) MITSUBISHI KASEI CORP.
 XX
 PI WPI: 1995-144713/19.
 PI P-PSDB: AAR72801.
 DR
 DR
 PT Antiviral agent comprising component which disrupts viral gene
 PT translation - used for the selective inhibition of e.g. Hepatitis
 PT C virus, polio: virus and human rhinovirus
 XX
 XX
 PS Disclosure; Pages 19-22; 23pp; Japanese.
 XX
 CC AA086798 encodes AAR72800 Hepatitis C virus (HCV) T7N1-19, which
 CC disrupts viral gene translation, by preventing the binding of
 CC the viral mRNA to the internal ribozyme entry site. It was
 CC used in the construction of an antiviral agent detecting vector
 CC (AA086799 which encodes AAR72801), which comprises a vaccinia virus
 CC vector, a vector containing T7N1-19 and a firefly luciferase gene.
 CC This antiviral agent can be used to treat HCV, poliovirus, cerebral
 CC cardio-hepatitis, human rhinovirus and foot and mouth disease viral
 CC infections.
 CC
 XX
 SQ Sequence 4987 BP; 1516 A; 1050 C; 1081 G; 1322 T; 18 other;

Query Match 95.6%; Score 326.8; DB 16; Length 4987;
 Best Local Similarity 99.1%; Pred. No. 2.1e-87;

PN WO9634976-A1.
 XX 07-NOV-1996.
 PD
 XX 01-MAY-1996; 96WO-US06070.
 PF
 XX 01-MAY-1995; 95US-0432693.
 PR
 XX (VERT-) VERTEX PHARM INC.
 PA
 XX Su M;
 PI
 XX WPI: 1996-506176/50.
 DR
 XX Assay for protease activity, esp. viral protease - using host cells
 PT which secrete exogenous or endogenous proteases and an artificial
 PP polypeptide substrates
 PS
 XX Example 1: Page 21-25; 44pp; English.
 CC A full-length hepatitis C virus (HCV) H strain cDNA (AAT41882)
 CC includes regions coding for the NS3 and NS4 proteases. NS3 is a
 CC 70 kDa protease involved in the maturation of viral polypeptides
 CC following infection. A novel method for assaying protease (e.g.
 CC NS3) activity involves transforming a host cell with DNA encoding
 CC the protease and DNA encoding a substrate (e.g. interleukin-1 beta).
 CC When these nucleotide sequences are expressed, the protease cleaves
 CC the substrate, which is secreted from the host where it can be
 CC easily quantitated. The method can be used to identify potential
 CC inhibitors of the protease for therapeutic appln.
 CC
 XX
 SQ Sequence 9401 BP; 1868 A; 2848 C; 2686 G; 1999 T; 0 other;

Query Match 95.1%; Score 325.2; DB 17; Length 9401;
 Best Local Similarity 98.8%; Pred. No. 7.3e-87;
 Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagccccctgattggtggcgagacaccccatatagatcctccctgtgaggaactactg 60
 DB 1 gccagccccctgattggtggcgagacaccccatatagatcctccctgtgaggaactactg 60
 QY 61 tcttcacgagaaagcgcttaagcattgagtgatgagtgctgagcctccagagcc 120
 DB 61 tcttcacgagaaagcgcttaagcattgagtgatgagtgctgagcctccagagcc 120
 QY 62 tcttcacgagaaagcgcttaagcattgagtgatgagtgctgagcctccagagcc 120
 DB 62 tcttcacgagaaagcgcttaagcattgagtgatgagtgctgagcctccagagcc 120
 QY 121 ccccccctccgggagagcattatgctgaggaacccgtgagtaacccggaattgcag 180
 DB 121 ccccccctccgggagagcattatgctgaggaacccgtgagtaacccggaattgcag 180
 QY 122 ccccccctccgggagagcattatgctgaggaacccgtgagtaacccggaattgcag 180
 DB 122 ccccccctccgggagagcattatgctgaggaacccgtgagtaacccggaattgcag 180
 QY 181 gacgagccgggtctcttcttctggaatcccgctcaatgctgagatttggcggtccccc 240
 DB 181 gacgagccgggtctcttcttctggaatcccgctcaatgctgagatttggcggtccccc 240
 QY 241 cgcgagactgctcagcagatgctggtgctgcgaaagccctgtgactcccgatag 300
 DB 241 cgcgagactgctcagcagatgctggtgctgcgaaagccctgtgactcccgatag 300
 QY 301 ggtgctgcgagtgcccgaggaggtctcgtagaccgtgcacc 342
 DB 301 ggtgctgcgagtgcccgaggaggtctcgtagaccgtgcacc 342
 QY 300 ggtgctgcgagtgcccgaggaggtctcgtagaccgtgcacc 341
 DB 300 ggtgctgcgagtgcccgaggaggtctcgtagaccgtgcacc 341

RESULT 9
 ID AAX24843 standard; DNA; 9595 BP;
 AAX24843;
 AC
 XX 21-JUN-1999 (first entry)
 DT
 XX Infectious hepatitis C virus genotype 1b strain HC-J4 genome.
 DE
 XX

KW HCV; Infectious clone; infection; diagnosis; therapy; vaccine;
 KW screening; assay; antiviral; virucide; ss.
 XX
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 ET CDS 342..9374
 FT CDS /*tag= a
 XX
 PN WO9904008-A2;
 XX
 PD 28-JAN-1999.
 XX
 PF 16-JUL-1996; 96WO-US14688.
 XX
 PR 27-JAN-1998; 98US-0014416.
 PR 18-JUL-1997; 97US-0053062.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Burk J, Emerson SU, Purcell RH, Yanagi M;
 XX
 P 1999-132252/11.
 P-PSDB: AAM98022.
 XX
 CC New isolated hepatitis C virus nucleic acids - used to develop
 CC products for the diagnosis, prevention and treatment of HCV
 CC infections and for developing screening assays
 CC
 XX
 PS Claim 3; Fig 14A-F; 126pp; English.
 XX
 CC The present sequence comprises the nucleic acid sequence of the
 CC genome of infectious hepatitis C virus (HCV) genotype 1b strain
 CC HC-J4 (ATCC 209596) that is capable of expressing this virus when
 CC transfected into cells. HC-J4 was obtained from acute phase plasma
 CC of a chimpanzee experimentally infected with serum containing
 CC HC-J4/91. The claimed infectious nucleic acid sequence can be used
 CC to produce chimeric genomes (see AAX24833) consisting of the open
 CC reading frames of infectious nucleic acid sequences of other
 CC genotypes (including genotypes 1-6) and subtypes (such as 1b, 2a,
 CC 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also relates to
 CC the introduction of mutations or deletions into infectious nucleic
 CC acid sequences in order to produce an attenuated HCV virus suitable
 CC for vaccine development. Infectious nucleic acid sequences can
 CC also be used to produce attenuated virus via passage in vitro or in
 CC vivo of the viruses produced by transfection of a host cell with
 CC the infectious nucleic acid sequence. Vaccines comprising one or
 CC more polypeptides made from the infectious nucleic acid sequence are
 CC used to immunise mammals, especially humans, against hepatitis C.
 CC The nucleic acid sequences can also be used to induce protective
 CC immunity against the virus. The nucleic acid sequences or their
 CC encoded proteases (e.g. NS3 protease) can additionally be used to
 CC develop screening assays to identify antiviral agents for HCV.
 CC
 XX
 SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;

Query Match 95.1%; Score 325.2; DB 20; Length 9595;
 Best Local Similarity 98.8%; Pred. No. 7.3e-87;
 Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagccccctgattggtggcgagacaccccatatagatcctccctgtgaggaactactg 60
 DB 1 gccagccccctgattggtggcgagacaccccatatagatcctccctgtgaggaactactg 60
 QY 61 tcttcacgagaaagcgcttaagcattgagtgatgagtgctgagcctccagagcc 120
 DB 61 tcttcacgagaaagcgcttaagcattgagtgatgagtgctgagcctccagagcc 120
 QY 121 ccccccctccgggagagcattatgctgaggaacccgtgagtaacccggaattgcag 180
 DB 121 ccccccctccgggagagcattatgctgaggaacccgtgagtaacccggaattgcag 180

QY 181 gagacacgggtccttcttcttgatcaatccgcgtcaatgcttgagatttgagcgtgcc 240
DB 181 gagacacgggtccttcttcttgatcaatccgcgtcaatgcttgagatttgagcgtgcc 239
QY 241 cgcgagactgctacgcgagtagtgtgtgtcgagaaaggccttggttactgctgtag 300
DB 240 cgcgagactgctacgcgagtagtgtgtgtcgagaaaggccttggttactgctgtag 299
QY 301 ggtgcttcgagtgccccggagaggtctcgttagaccgtgcacc 342
DB 300 ggtgcttcgagtgccccggagaggtctcgttagaccgtgcacc 341
RESULT 10
AAC86939 standard; DNA; 9595 BP.
ID AAC86939 standard; DNA; 9595 BP.
XX AAC86939;
AC AAC86939;
XX 02-APR-2001 (first entry)
DT 02-APR-2001 (first entry)
DE Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 1b.
XX Chimeric virus: bovine viral diarrhoea virus; BVDV; hepatitis C virus;
KM HCV; vaccine; viral inhibitor; antiviral; ss.
XX Hepatitis C virus.
OS Hepatitis C virus.
XX Key Location/Qualifiers
FH CDS 342..9374
FT /*tag= a
XX WO200075352-A2.
PN 14-DEC-2000.
PD 02-JUN-2000; 2000MO-US15527.
XX 04-JUN-1999; 9905-0137817.
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Nam J, Bukh J, Emerson SU, Purcell RH;
XX WPI; 2001-071081/08.
DR P-PSDB; AAB31170.
XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus
PT genome in which the (non-)structural region has been replaced by
PT hepatitis C virus (HCV) genome useful for treating or preventing HCV
PT signs and symptoms
XX Disclosure; Fig 4A-F; 97pp; English.
XX The specification describes a nucleic acid comprising a chimeric virus
CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
CC the (non-)structural region has been replaced by the (non-)structural
CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
CC the chimeric virus and the chimeric virus are useful for identifying
CC cell lines capable of supporting the replication of these chimeric
CC viruses, in screening for neutralizing antibodies to HCV of different
CC genotypes, in the production of HCV-BVDV viruses, for the development
CC of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal,
CC in studying the molecular properties of HCV indirectly in vitro, and in
CC identifying inhibitors of viral enzyme activity which would be useful
CC as antiviral agents. Formulations or compositions comprising the
CC chimeric viruses may be used to treat or prevent the signs and symptoms
CC of HCV. The present sequence represents a HCV clone, which is used
CC to construct chimeric nucleic acids of the invention.
XX Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;
SQ

Query Match 95.18; Score 325.2; DB 22; Length 9595;
Best Local Similarity 98.88; Pred. No. 7.3e-87;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 ggcagccccctgattggggcgagacatccacatgatgactccctcttgagaaactctg 60
DB 1 ggcagccccctgattggggcgagacatccacatgatgactccctcttgagaaactctg 60
QY 61 tcttcacgacgaagaagcgtctctagccatgagcttatgagtgtctgacgctcagagcc 120
DB 61 tcttcacgacgaagaagcgtctctagccatgagcttatgagtgtctgacgctcagagcc 120
QY 121 ccccccctcccgaggagacatagtggtctcgagaaaggccttggttactgctgtag 180
DB 121 ccccccctcccgaggagacatagtggtctcgagaaaggccttggttactgctgtag 180
QY 181 gagacacgggtccttcttcttgatcaatccgcgtcaatgcttgagatttgagcgtgcc 240
DB 181 gagacacgggtccttcttcttgatcaatccgcgtcaatgcttgagatttgagcgtgcc 239
QY 241 cgcgagactgctacgcgagtagtgtgtgtcgagaaaggccttggttactgctgtag 300
DB 240 cgcgagactgctacgcgagtagtgtgtgtcgagaaaggccttggttactgctgtag 299
QY 301 ggtgcttcgagtgccccggagaggtctcgttagaccgtgcacc 342
DB 300 ggtgcttcgagtgccccggagaggtctcgttagaccgtgcacc 341
RESULT 11
AAF23492
ID AAF23492 standard; DNA; 9595 BP.
XX AAF23492;
AC AAF23492;
XX 21-MAR-2001 (first entry)
DT 21-MAR-2001 (first entry)
DE Infectious Hepatitis C virus 1b genotype.
XX Infectious Hepatitis C virus 1b genotype.
XX GRV-B; hepatitis C virus; HCV; vaccine; ds.
OS Hepatitis C virus.
XX KC200075337-A1.
FN 14-DEC-2000.
PD 02-JUN-2000; 2000MO-US15293.
XX 04-JUN-1999; 9905-0137694.
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX WPI; 2001-091214/10.
DR New infectious nucleic acids of the GB virus-B clone, useful for
PT indirectly studying the molecular properties of hepatitis C virus (HCV)
PT and in developing vaccines and therapeutics for HCV
XX Disclosure; Fig 7; 96pp; English.
XX The present invention relates to GB virus-B. The nucleic acid molecules
CC of the invention are useful for indirectly studying the molecular
CC properties of hepatitis C virus (HCV). The infectious nucleic acid
CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC in the development of vaccines and therapeutics for HCV.
XX Sequence 9595 BP; 1934 A; 2843 C; 2697 G; 2121 T; 0 other;
SQ
Query Match 95.18; Score 325.2; DB 22; Length 9595;

Best Local Similarity 98.8%; Pred. No. 7.3e-87;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagcccccgtatggggcgacacccacacatagatccctccctgtggaactacty 60
DB 1 gccagcccccgtatggggcgacacccacacatagatccctccctgtggaactacty 60
QY 61 tcttcacgagaagaacgctatgacatgacgttagatgagtgctgtcagccctcagcc 120
DB 61 tcttcacgagaagaacgctatgacatgacgttagatgagtgctgtcagccctcagcc 120
QY 121 ccccccctccgggagagacatagtgctcgcggaaccggctgaaatcacccgaaatccag 180
DB 121 ccccccctccgggagagacatagtgctcgcggaaccggctgaaatcacccgaaatccag 180
QY 181 gacgacggggtcccttctctgtgacaaatcccgctcaaatgtccttgagaaatttggtcgtcc 240
DB 181 gacgacggggtcccttctctgtgacaaatcccgctcaaatgtccttgagaaatttggtcgtcc 240
QY 241 cgcgagactgctagaccgaatagtgctggtcgcggaagccctgtgtactcctcgtatg 300
DB 241 cgcgagactgctagaccgaatagtgctggtcgcggaagccctgtgtactcctcgtatg 300
QY 301 ggtgcttgcaagtgcccgagagctcgtctgtagaccgtgacc 342
DB 301 ggtgcttgcaagtgcccgagagctcgtctgtagaccgtgacc 341

RESULT 12

AAK24833
ID AAK24833 standard; DNA; 9599 BP.

AC AAK24833;

DT 21-JUN-1999 (first entry)

DE Infectious hepatitis C virus genotype 1a/1b chimera genome.

KM HCV; infectious clone; infection; diagnosis; therapy; vaccine;

KM screening; assay; antiviral; virucide; ss.

OS Hepatitis C virus.

FH Key Location/Qualifiers

FT CDS 342..9377

WT WO9904008-A2.

PD 28-JAN-1999.

PF 16-JUL-1998; 98WO-US14688.

PR 27-JAN-1998; 98US-0014416.

PR 18-JUL-1997; 97US-0053062.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Burk J, Emerson SU, Purcell RH, Yanagi M;

DR WPI: 1999-132252/11.

DR P-PSDB; AAM98021.

XX New isolated hepatitis C virus nucleic acids - used to develop

XX products for the diagnosis, prevention and treatment of HCV

XX infections and for developing screening assays

XX Claim 8; Fig 16A-F; 126pp; English.

XX This nucleic acid sequence comprises the genome of infectious

XX hepatitis C virus (HCV) chimeric 1a/1b clone pH7C-J4 which

XX contains the nonstructural region of infectious genotype 1a strain

XX H77 (see AAK24832) and the structural region of infectious genotype

CC 1b strain HC-14 (see AAK24833). The construction of such chimeric
CC nucleic acid sequences is expected to be of importance in studying
CC the growth and virulence properties of HCV and in the production
CC of HCV suitable for conferring protection against multiple
CC genotypes of HCV. The invention also relates to the introduction
CC of mutations or deletions into infectious nucleic acid sequences
CC in order to produce an attenuated HCV virus suitable for vaccine
CC development. Infectious nucleic acid sequences can also be used to
CC produce attenuated virus via passage in vitro or in vivo of the
CC viruses produced by transfection of a host cell with the infectious
CC nucleic acid sequence. Vaccines comprising one or more
CC polypeptides made from the infectious nucleic acid sequence are
CC used to immunise mammals, especially humans, against hepatitis C.
CC The nucleic acid sequences can also be used to induce protective
CC immunity against the virus. The nucleic acid sequences or their
CC encoded proteases (e.g. NS3 protease) can additionally be used to
CC develop screening assays to identify antiviral agents for HCV.

SD Sequence 9599 BP; 1889 A; 2873 C; 2724 G; 2113 T; 0 other:

Query Match 95.1%; Score 325.2; DB 20; Length 9599;

Best Local Similarity 98.8%; Pred. No. 7.3e-87;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagcccccgtatggggcgacacccacacatagatccctccctgtggaactacty 60
DB 1 gccagcccccgtatggggcgacacccacacatagatccctccctgtggaactacty 60
QY 61 tcttcacgagaagaacgctatgacatgacgttagatgagtgctgtcagccctcagcc 120
DB 61 tcttcacgagaagaacgctatgacatgacgttagatgagtgctgtcagccctcagcc 120
QY 121 ccccccctccgggagagacatagtgctcgcggaaccggctgaaatcacccgaaatccag 180
DB 121 ccccccctccgggagagacatagtgctcgcggaaccggctgaaatcacccgaaatccag 180
QY 181 gacgacggggtcccttctctgtgacaaatcccgctcaaatgtccttgagaaatttggtcgtcc 240
DB 181 gacgacggggtcccttctctgtgacaaatcccgctcaaatgtccttgagaaatttggtcgtcc 240
QY 241 cgcgagactgctagaccgaatagtgctggtcgcggaagccctgtgtactcctcgtatg 300
DB 241 cgcgagactgctagaccgaatagtgctggtcgcggaagccctgtgtactcctcgtatg 300
QY 301 ggtgcttgcaagtgcccgagagctcgtctgtagaccgtgacc 342
DB 301 ggtgcttgcaagtgcccgagagctcgtctgtagaccgtgacc 341

RESULT 13

AAO32436
ID AAO32436 standard; DNA; 7911 BP.

AC AAO32436;

DT 26-APR-1993 (first entry)

DE HCV antigen clone T7N1-30.

KM Clone; Hepatitis C virus; HCV; core-envelope; NS1(gp70); NS2-NS4;

KM NS4-NS5; region; diagnostic method; antibody; suppress; control;

KM proteolytic; process; precursor; polypeptide; ss.

OS Hepatitis C virus.

FH Key Location/Qualifiers

FT CDS 369..7898

WT BP518313-A.

PD 16-DEC-1992.


```

DB 181 gagcagcgggtctctcttctggtatcaa-ccgcctcaatgctctggaatcttgaggcgtgcccc 239
OY 241 cgcgagacgtctagaccgagtagtgtggtgcgaaaggccttctgtagctctgtagatag 300
DB 240 cgcgaagacgtctagaccgagtagtgtggtgcgaaaggccttctgtagctctgtagatag 239
OY 301 ggtgctctgagatgcccgcggaggtctctgtagaccgtgacc 342
DB 300 ggtgctctgagatgcccgcggaggtctctgtagaccgtgacc 341

RESULT 15
AAZ08979 standard; DNA; 341 BP.
AAZ08979:
02-NOV-1999 (first entry)
Hepatitis C virus 5' UTR used as promoter for a HCV genetic vaccine.
Genetic vaccine; immunisation; humoral response; cellular response;
non-structural protein; NS protein; antigen; ds.
Hepatitis C virus.
W09938880-A1.
05-AUG-1999.
28-JAN-1999; 99WO-US01823.
30-JAN-1998; 98US-0073156.
(GEHO ) GEN HOSPITAL CORP.
Encke J, Wands J;
WPI: 1999-494077/41.
New hepatitis virus nucleic acids for, e.g. inducing an immune
response against the virus
Claim 6: Page 10-11; 41pp; English.
This sequence represents the hepatitis C virus (HCV) 5' untranslated
region (5' UTR). This is preferred for use as a promoter in a
recombinant DNA vaccine against hepatitis C. The genetic
vaccine comprises regulatory elements capable of directing
expression in humans and one or more HCV non-structural
(NS) proteins. The HCV genome encodes three NS proteins: NS3, a serine
protease/viral helicase; NS4; and NS5, the viral RNA-dependent RNA
polymerase. In HCV infection, these and other virally encoded proteins
are produced by proteolytic processing of a precursor polypeptide.
However, for use in a genetic vaccine, the sequences coding for NS
proteins had to be engineered so that each NS sequence had its own
initiation and stop codons. Restriction sites were also engineered
into the gene fragments to aid subcloning. The gene fragments were
isolated and mutated using PCR. The NS3 coding sequence was modified
using PCR primers AAZ08980 and AAZ08981, the NS4 coding sequence
modified using AAZ08984 and AAZ08985, and the NS5 coding sequence
modified using AAZ08986 and AAZ08987. Studies in mice immunised with
constructs expressing one NS protein were found to elicit strong
antigen-specific immune responses in both arms of the immune system. This
demonstrated that the NS proteins are better antigens for stimulating
humoral immune responses as compared with previous studies using the HCV
core structural protein. The genetic vaccine is useful for inducing
an immune response (cellular or humoral) against hepatitis C virus in
a human uninfected by the virus, and for immunising a human susceptible
to hepatitis C viral infection by inducing an immune response. The
composition is also useful for treating a human infected with hepatitis
C virus, by induction of an immune response. The advantage of this

```

```

CC method of immunisation compared with immunisations with soluble
CC recombinant proteins or peptides, is the ability to induce a strong
CC inflammatory CD4+ T cell response as well as cytotoxic T cell activity.
CC In addition, the new recombinant genetic vaccine is more suitable for
CC immunisation; unlike synthetic peptides which only have a limited
CC number of epitopes available for stimulation of the host response.
SQ Sequence 341 BP; 63 A; 103 C; 106 G; 69 T; 0 other;
Query/Match 94.6%; Score 323.6; DB 20; Length 341;
Best Local Similarity 98.5%; Pred No. 9,4e-87;
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
OY 1 gccagcccccctgtaggggagacactccacatagatcccccctgtgaggaactactg 60
DB 1 gccagcccccctgtaggggagacactccacatagatcccccctgtgaggaactactg 60
OY 61 tcttcacgagaagagcgtctagcagatggttagttagtgcgtgcagcctcagacc 120
DB 61 tcttcacgagaagagcgtctagcagatggttagttagtgcgtgcagcctcagacc 120
OY 121 ccccccctccggagagacatagtgctgcggaacggtagtacccggaattgccag 180
DB 121 ccccccctccggagagacatagtgctgcggaacggtagtacccggaattgccag 180
OY 181 gagcagcgggtcccttctctgtagtaatccgcgtcctgaagatttggtggtgcccc 240
DB 181 gagcagcgggtcccttctctgtagtaatccgcgtcctgaagatttggtggtgcccc 240
OY 241 cgcgagacgtctagaccgagtagtgtggtgcgaaaggccttctgtagctctgtagatag 300
DB 240 cgcgagacgtctagaccgagtagtgtggtgcgaaaggccttctgtagctctgtagatag 299
OY 301 ggtgctctgagatgcccgcggaggtctctgtagaccgtgacc 342
DB 300 ggtgctctgagatgcccgcggaggtctctgtagaccgtgacc 341

```

Search completed: March 1, 2002, 15:30:43
Job time: 25051 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 11:35:41 ; Search time 8196.34 Seconds

(without alignments)
448,378 Million cell updates/sec

Title: US-09-763-836-7

Perfect score: 342

Sequence: 1 gccagcccccgtatg99g9c.....ggtctcgtagacgltgcaac 342

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 segs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gp_est1:*
11: gp_est2:*
12: gp_hic:*
13: gp_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40.8	11.9	925	13	CNS0091P
2	38	11.1	925	13	CNS0091P
3	35.6	10.4	844	13	CNS0052P
4	35.2	10.3	600	13	AO338955
5	34.2	10.0	548	10	AW958779
6	34.2	10.0	610	10	AV654578
7	34	9.9	394	11	BF901361
8	34	9.9	440	11	BE845986
9	33.4	9.8	421	10	AV397071
10	33.4	9.8	432	10	AA196987
11	33.4	9.8	777	11	BF859305
12	32.8	9.6	371	11	BF508846

13	32.8	9.6	405	13	AO815897
14	32.8	9.6	849	11	BE759010
15	32.8	9.6	927	11	BG492855
16	32.8	9.6	1101	13	CNS0052P
17	32.6	9.5	645	13	CNS01213
18	32.6	9.5	952	10	AL563471
19	32.4	9.5	339	10	AV686380
20	32.4	9.5	369	10	AV693842
21	32.4	9.5	370	10	AV693850
22	32.4	9.5	575	11	BF001581
23	32.4	9.5	577	13	AZ224104
24	32.4	9.5	961	13	CNS01087
25	32.2	9.4	400	10	BE498161
26	32.2	9.4	470	11	BG263335
27	32.2	9.4	977	13	CNS0091P
28	32.2	9.4	430	11	HS7082
29	32	9.4	897	11	BG031227
30	31.8	9.3	436	10	AI094710
31	31.6	9.3	778	10	BE131094
32	31.6	9.2	530	11	BE349419
33	31.6	9.2	793	10	BE408210
34	31.4	9.2	638	13	AZ003859
35	31.4	9.2	901	11	BI414151
36	31.2	9.1	949	11	BE799450
37	31.2	9.1	442	10	BE728496
38	31.2	9.1	549	10	BE424949
39	31.2	9.1	844	13	CNS0052P
40	31.2	9.1	1003	10	AL523533
41	31.2	9.1	1006	11	BI104877
42	31	9.1	387	10	AA921613
43	31	9.1	647	10	AV650450
44	31	9.1	648	10	AV650417
45	31	9.1	683	13	AZ125374

ALIGNMENTS

RESULT 1	CNS0091P	925 bp	DNA	GSS	03-JUN-1999
LOCUS	CNS0091P				
DEFINITION	Drosophila melanogaster genome survey sequence TERT end of BAC # BACR19D16 of RPT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL053013	GI:4934461			
VERSION	AL053013				
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 925)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila genome project (BDGP).
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES		Location/Qualifiers
source	1..925	/organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="RPCI-98" /clone="BACR19D16" /note="end : TERT3"
BASE COUNT ORIGIN	120 a 61 c 61 g 172 t 511 others	
Query Match	11.9% ; Score 40.8 ; DB 13; Length 925;	
Best Local Similarity	14.9%; Pred. No. 0.28;	
Matches	Conservative 43; Mismatches 117; Indels 0; Gaps 0	
OY	48 ttgaagaactactgtcttcacgcagaaagcglctaggcatlgygcltagtatlgaatlgatcgltg 107	
DB	502 TTATTAANNNNNANNNANNNNNAGCSCMSCKKCGSTTBGTTTTTSSSGSYG 561	
OY	108 cagcccccagagcccccccctccggagagcatagttgtctgcggaaaccgltagtataa 167	
DB	562 KCCSSGSGBSCCSCSCSCSSSCCBCCCSCCSSTCCSSBSBKCSSTSBCSCCC 621	
OY	166 ccggaattgccagacgcagcgtctctcttttgatcaalccgcgtccaatgcctgatgat 227	
DB	622 SSKVCCTGCSSSSSSSSSTSSSTSSTSSKSSSSSSSSSSSSTTTTKSTAASGG 681	
OY	228 ttgggcgtgccccgcgagacctgctagccgagtagtgtgtggatgcgcaaaggcctgtgt 287	
DB	682 SMSAGGSGSTSTSSSSSSSTSSSVSGSMSTBSGSBSBGSSSSSSSTSBBS 741	
OY	288 tacttcctatatagggtcgtctgcagatgccccggagagctcgttaaacg 336	
DB	742 CTSTSSSSSSSTCCTCCTCCCTSYSTSSSTSSSTSWGSTSGSSS 790	
RESULT 2	CNS0091P/C	GSS 03-JUN-1999
LOCUS	CNS0091P	925 bp DNA
DEFINITION	Drosophila melanogaster genome survey sequence TERT3 end of BAC # BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
ACCESSION	AL053013	
VERSION	AL053013.1 GI:4934461	
KEYWORDS	GSS.	
SOURCE ORGANISM	Fruit fly. Drosophila melanogaster Euarystia; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	Genoscope.	
AUTHORS TITLE JOURNAL COMMENT	Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACS. For further information please see http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.	
FEATURES source	Location/Qualifiers 1..925	

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RP11-98"
/clone="BACRP19D16"
/notes="end : TET3"

BASE COLUMN      120 a      61 c      61 g      172 t      511 others
ORIGIN

Query Match
Best Local Similarity 11.1%; Score 38; DB 13; Length 925;
Matches 35; Conservative 117; Mismatches 104; Indels 0; Gaps 0;

Cy      69 cagaagcgtctgacgcatgctgtgtatgtagtgcgtgcgaacccctccagcccccctc 128
Cb      829 CGSASARGVKVMASGAGAKRGSGGAGASMSHSSSSACBSSSSSCASCMSASSSSAS 770
Cy      139 ccggagagagcattatgtctctgcgaaccggtgtatcacacggaaattgcaggacgac 188
Db      769 RRRSCGAGAGSASSSSSSSSSSSASAGSVSSASSSSSSSSSVSCSVASMSCSBPS 710
Cy      189 agtccctcttgatcaatcccccgaatgcttgagattgtggcgtgccccgcgcgac 248
Db      709 SSSASASSSSSSSSASCSACSCCCTTSMSCCTTSASMSARSSSSSSSCSSSMAS 650
Cy      249 tgcctgcgagatgagtgtgttggtgcgaaggcctctgttactgcctatgagtgtctg 308
Db      649 ASSSSASSSSSSSSSSGCSACGBSMSSGCGSGSVSSASAGMSVSSSGRRSGSGGG 590
Cy      309 cgagtcgcccggaagg 324
Db      589 VCGSSSSSSSGSSSGC 574

RESULT : 3
LOCUS      CNS0052P
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACRP19D16 of RP11-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL056652
VERSION     AL056652.1 GI:4932342
KEYWORDS    GSS.
SOURCE      fruit fly.
            Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Plekaryota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
ORGANISM    Genoscope.
            Direct Submission
            Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
JOURNAL     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mamoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RP11-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain Y2; cn bw sp. the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
            Location/Qualifiers
            1..844
FEATURES
SOURCE

```

BASE COUNT	261 a	112 c	92 g	35 t	344 others
ORIGIN					
Query Match	10.4%	Score 35.6;	DB 13;	Length 844;	
Best Local Similarity	15.7%	Pred. No. 8;			
Matches 37; Conservative 106; Mismatches 92; Indels 0; Gaps 0;					
Oy	104	cgtagagcctcagagccccccctcccgaggagagcatagtgctgaggaacgggtgag	163		
Db	403	SRSRMSSSCSSCAGSSCGVSSSCSAMSSSCCCVSVSCGAMSSCCSGGMSASSSSSSG	462		
Oy	164	tacacgggaatgcccaagagacgcgcggtcttcttgatcaatccgcctcaatgcctg	223		
Db	463	VSSVSGRAGVAGRGCMCSCMVOCCMCSSMCCSVCACVCSGVSRAVAGCSVGG	522		
Oy	224	agattgggcgtgcccccgagagactgctagcgaagtgtgtgtggtgcgaagacct	283		
Db	523	RVGGSRRRAGRSSRGSGSSVSSSVSSSSVCGMCACSSASVSCSSBSVASSVSG	582		
Oy	284	gtgagacgcctcgaatgaagggtctgtgcggtgctcccgaggaggtcctgtgacgctg	338		
Db	583	VSRCGRCVGGAGVGGSSRVSSSCSSSSSSSSGCGSGVSRSRGSAGVAVGSG	637		
RESULT 4					
A0538955/c	A0538955	600 bp	DNA	GSS	19-MAY-1999
LOCUS	RPCI-11-343B22.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-343B22				
DEFINITION	A0538955, DNA sequence.				
ACCESSION	A0538955				
VERSION	A0538955.1 GI:4868938				
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and Venter,J.C.				
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building				
JOURNAL	Unpublished (1997)				
COMMENT	other GSSs: RPCI-11-343B22.TV Contact: Shaying Zhao, William Niernan, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbeetlgr.org Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@eijong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tlgr.org/tld/humgen/bac_end_search/bac_end_search.html. Seq primer: SP6 Class: BAC ends.				
FEATURES	Location/Qualifiers				
source	1..600				
	/organism="Homo sapiens"				
	/db_xref="GDB:763137"				
	/db_xref="taxon:9606"				
	/clone="RPCI-11-343B22"				
	/clone_lib="RPCI-11"				
	/sex="Male"				
	/cell_type="Lymphocytes"				
	/note="vector: pBAC6.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"				

BASE COUNT	129 a	155 c	151 g	165 t
ORIGIN				
Query_Match	10.3%	Score 35.2;	DB 13;	Length 600;
Best Local Similarity	53.7%	Pred. No. 9.5;		
Matches	73;	Conservative 0;	Mismatches 63;	Indels 0;
Db	332	GAGAGCAATAGTGTATCGGAGCACACTGGAGGTACAACTTATGCTTGAGAGGCTCGCA	323	
Qy	193	cttcttgcatacaatcccgctcaatgctgtgagatttggcggtgccccgagactgct	252	
Db	332	CTGTGAGGAGAACCCACGATCAATTGTGTATATGTTGACACCTTCAGGAGACTGCT	263	
Qy	253	agccgagtagtggtg	268	
Db	252	AGCCGTACTGGGGTGG	247	
RESULT	5			
LOCUS				
AW958779	548 bp	mRNA	EST	01-JUN-2000
DEFINITION	ES3370849	MAGE resequences, MAGE Homo sapiens	CDNA, mRNA sequence.	
ACCESSION	AW958779			
VERSION	AW958779.1	GI:8148463		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 548)			
AUTHORS	Heide, P., Qi, R., Abernathy, K., Dharp, S., Gaspard, R., Gay, C., Holt, J.E., Seede, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.			
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: John Quackenbush The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johngetlgr.org Plate: 125 Seq primer: Reverse.			
FEATURES	Seq primer: Reverse.			
SOURCE	Location/Qualifiers			
1..548	/organism="Homo sapiens"			
/db_xref="taxon:9606"				
/clone_lib="MAGE resequences, MAGE"				
/note="Vector: pBluescriptsM"				
BASE COUNT	111 a	161 c	149 g	127 t
ORIGIN				
Query_Match	10.0%	Score 34.2;	DB 10;	Length 548;
Best Local Similarity	50.3%	Pred. No. 18;		
Matches	84;	Conservative 0;	Mismatches 83;	Indels 0;
Db	8	ccctgatgaggcgacatccaccatagatcaatccctcgtlagaagactactgtctc	67	
Qy	155	CTCAGATATTGAGGGGCTCTGGAAGCCTAGTCTCTCTTCTCTGACACAGCTGTGGCTGCC	214	
Db	98	gcagaagcgcgtctagcatgagcgtctagatagtgctgacgacctcagccccct	127	
Qy	215	CGTGCTCTCTTGATGACATGATTTAGCCCTACGTGGGCTGCAGCATTTGGGATTCAGGCT	274	
Db	138	ccgggagagccatagtgctgcggaacccggtgtgtacacccgagat	174	
Qy	275	ACCTAGAGGGGATCGGGCCAGGAAACCTTCGATTTAGCAAGCAAT	321	

RESULT	6	AV654578	610 bp	MRNA	EST	07-SEP-2000
LOCUS	AV654578	G1C Homo sapiens cDNA clone G1CDXC10.3			3'	mRNA sequence.
DEFINITION	AV654578					
ACCESSION	AV654578					
VERSION	AV654578.1	GI:9875592				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
AUTHORS	1 (bases 1 to 610)					
	Qian,B., Wu,T., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.					
	Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu					
	,G., Tang,Y., Gu,Y., Chen,Z., and Han,Z.					
TITLE	Homo sapiens cDNA clone					
JOURNAL	Unpublished (2000)					
COMMENT	Contact: Zeguan Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzq@chgc.sh.cn This clone is available at CHGC in Shanghai.					
FEATURES	Location/Qualifiers					
Source	1..610					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone="G1CDXC10"					
	/clone_1lb="G1C"					
	/tissue_type="corresponding non cancerous liver tissue"					
	/dev_stage="Adult"					
	/lab_host="SOLR"					
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"					
BASE COUNT	118 a 186 c 184 g 121 t 1 others					
ORIGIN						
Query Match	10.0%; Score 34.2; DB 10; Length 610;					
Best Local Similarity	50.3%; Pred. NO. 18;					
Matches	84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;					
QY	8 cccctaatgaggcgacatccacccatagatcactcccctgtgaggaactactgtcttcac 67					
DB	317 CTCACATATTGAGGGGCTCGAAGCCCTAGTTCGTCTCTCTCTGACAGCTGTGGCTTCCC 376					
OY	68 gacgaagagcgctctagatccatgagcgcttagatgtgtgtgagagcctccagagccccccct 127					
DB	377 CCGTGGCTGCTTGCTGACATGCAATTTAGCGCTACGTGGCGCTGCACGCAATTTGGGATTCACAGGCT 436					
OY	128 cccgggagagccatagtggtgtcgcggaacccggtgagttacacccgaat 174					
DB	437 ACCCTGAGGCGGCGCATCGGCGCCAGGAAACCTCGGATTTACACAGCAAT 483					
RESULT	7					
LOCUS	BF901361	394 bp	MRNA	EST	18-JAN-2001	
DEFINITION	PM2-MT0199-081200-001-a10 MT0199 Homo sapiens cDNA, mRNA sequence.					
ACCESSION	BF901361					
VERSION	BF901361.1	GI:12292820				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
REFERENCE	1 (bases 1 to 394)					
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordino,S., Costa,F.F.,					

TITLE	Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.					
JOURNAL	Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare					
MEDLINE	'M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and					
COMMENT:	Simpson,A.U.J. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags Proc. Natl. Acad. sci. U.S.A. 97 (7), 3491-3496 (2000) 20202663					
	Contact: Simpson A.U.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SF, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM2&t2=PM2-MT0199 081200-.001-a10&t3=2000-12-08&tc=1) Seq primer: puc 18 forward High quality sequence start: 18 High quality sequence stop: 174. Location/Qualifiers 1..394 /organism="Homo sapiens" /db_xref="taxon:9606" /cclone_lib="MT0199" /dev_stage="Adult" /note="Organ: marrow; Vector: puc18; Site:1: Sma1; Site:2: Sma1; A mini-library was made by cloning products derived from ONESRES PCR (U.S. Letters Patent application No. 199 , 716 - Ludwig Institute for Cancer Research) profiles into the PUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."					
BASE COUNT	92 a	124 c	116 g	72 t		
ORIGIN						
Query Match	9.9%; Score 34; DB 11; Length 394;					
Best Local Similarity	50.6%; Pred. NO. 18;					
Matches	82; Conservative	0; Mismatches	80; Indels	0; Gaps	0;	
Gy	120	ccccccctccgggagagcattagtgtctcggaacgcggtagataaccggaattgcc	179			
Dd	177	cctccgcagccccttccgcgccCGAGGACAGCAGCCGCCCGAAGAATACACGGCCAGGCMA	118			
Gy	180	gagacgacgcggtgcctttcttgatcaaacccgctaagcgcgtgatattgggtgcc	239			
Dd	117	ttagacctcttcAGCACACAGGTCTCCCATTCAGCATCCCTCCAGCACGCCCGGCTTCACCG	58			
Gy	210	ccgcgagactgtacgcagagtagtgttggttcgcgaaggcc	281			
Dd	57	tagagcaactgcattcgctctggccttgagcacagac	16			
RESULT#	8					
HEB45985						
LOCUS	BBE45986 440 bp EST 25-SEP-2000					
DEFINITION	232495 BANC 5BOV Bos taurus CDNA 5', mRNA sequence.					
ACCESSION	BBE45986					
VERSION	BBE45986.1 GI:10282801					
KEYWORDS	EST.					
SOURCE	cow.					
ORGANISM	Bos taurus					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos. 1 (bases 1 to 440) Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and Wells,K.D. Mapping of Expressed Sequence Tags from a normalized bovine mammary					

gland cDNA library
unpublished (2000)
Contact: Sonstegard TS

USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200, Rm 2A, Beltsville, MD 20705, USA
Tel.: 301 504 8416
Fax: 301 504 8414
Email: tads@lps1.barc.usda.gov
Single pass sequencing. Bases called and all-trimmed with Phred
v0.980904.e. Vector identified by cross-match with the -mscore 18
and -mismatch 12 options.

1. .440

```

/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_id="BARC_5BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6. Site_1: XbaI, Site_2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

```

90 a 135 c 133 g 82 t

Query Match	9.9%	Score 34;	DB 11;	Length 440;
Best Local Similarity	50.6%;	Pred. No. 19;		
Matches 82;	Conservative	0;	Mismatches 80;	Indels 0;
			Gaps	0;

0Y	59	tgcttcaacgaagaaacgctctagccatgagcgttatgatgatgtctcgttgacgctccagg	118
Db	189	TGACTTTAAGCACAACACTCCCGTCCACACAGCAGTTCCGTAATTACAGCCGCCATGGG	248
0Y	119	ccccccctcccgagagacatagtgctcggagaaacgctgagctacacgcgaattgac	178
Db	249	CCACCCTTCATGCTGCAGATACCTTTATTCCTTGAAAGGCTGTCATTAAGGACGACTC	308
0Y	179	agaaagacagggctccttctctgtatcaatcccgctcaatgac	220
Db	309	AGCCACAGCGGGTCTCTTCCTGGCTCTCTCCGCTGCATGCC	350

LOCUS	421 bp	EST	29-SEP-2000
AV397071	Chlamydomonas reinhardtii C9	Chlamydomonas reinhardtii	
AV397071	Chlamydomonas reinhardtii C9	Chlamydomonas reinhardtii	
CDNA clone CL75b06_r	mRNA sequence.		

AV39/U/1
AV397071.1 GI:6551287
EST.
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii

1 (bases 1 to 421)
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green

DNA Res. 6 (6), 369-373 (1999)
20152988
Contact: Yasukazu Nakamura

Email: ynakamu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/location/Qualifiers>

```
1, .421
/organism="Chlamydomonas reinhardtii"
/strain="C9"
```

```

/ab-xref="taxon:3055"
/clone="CL7506_x"
/clone_id="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescriptipti SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      83 a      147 c      114 g      77 t
ORIGIN

```

83 a 147 c 114 g 77 t

```

Query: Match 9.88; Score 33.4; DB 10; Length 421;
      Best Local Similarity 55.78; Pred. NO. 28;
      Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0

```

Qy .13 gtacccggaattccagagacacgcggtctctcttgatgacaaatccgcgtcaatgctg 222
Db .16 GCACACACACACACAGCTGTGTGGGGCGGTATTGTGGTGAACCAAGCCGACGCGCTG 57
Qy .23 gagatttggcgctgccccgcgcgagactgtctagccgagatagtglttgggtctgcgaaa 277
Db .56 GATGCTGGGAGCGAGACCGCTGGGCCCTGTGGCCCTACCTGTAAACGGAGACCGCA 2

AA196987	432 bp	mRNA	EST	22-JAN-1997
----------	--------	------	-----	-------------

DEFINITION
cg60h12.r1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
clone IMAGE:646055 5' similar to gb:J03040 SPARC PRECURSOR (HUMAN)
); mRNA sequence.
1106007

5987
5987.1 GI:1792578
n.
sapiens

Boases 1 to 432)

ier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Chisue, S., Dietrich, N., Dubnue, T., Favvello, A., Gish, N., Hawkins, M., Hultman, M., Kucaba, T., Laczy, M., Le, M., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Watson, R., Trevisakis, P., Underwood, K., Wohlmann, P., Yatsenfon, R., Wilson, R., and Matra, M. Generation and analysis of 250,000 human expressed sequence tags Genome Res. 6 (9), 807-822 (1996)

me Res. 6 (9), 807-828 (1996)
4478
act: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estwatson.wustl.edu
This clone is available royalty-free through INL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 227.

Location/Qualifiers

```

source
1. 432
/organism="Homo sapiens"
/db_xref="GDB:5216306"
/db_xref="taxon:9606"
/clone="IMAGE:646055"
/clone_1ib="Stratagene neuroepithelium (#937231)"
/seq_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/notes="vector: pBluescript SK-; Site-1: EcoRI; Site-2: XhoI; Cloned unidirectionally. Primer: Oligo dT. NT2 cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24 hours. Average insert size: 1.5 kb; Uni-TAP XR Vector; -5 adaptor sequence: 5' GAATTGGCAGCAG 3' -3' adaptor

```


Db 687 GAGAGCCCCCTGCTTGGACTATCCCTTCCATCCCTGTAGAGGAGAGAGGCG 632

RESULT 15
PCA93B55 /

BC6492855/c
LOCUS

DEFINITION 602536884F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4655685 5'

mrna sequence.
BGA92855

VERSION BG492855.1 GI:13454367

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 927)
Mammalia; Eulheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS
NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, N

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg Ph D

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 15:26:13 ; Search time 115.48 Seconds
(without alignments)
670.726 Million cell updates/sec

```

Title: US-09-763-836-7
Perfect score: 342
Sequence: 1 ggcagcccccctgattgggc.....ggtctcgtagaccgtgcacc 342

```

Scoring table:	IDENTITY_NUC
Canon 10 0	Canext 1 0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:*

```

1: /cgn2_6/ptodata/2/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/lna/PCtUS_COMB.seq: *
6: /cgn2_6/ptodata/2/lna/PCtUS1.seq: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3.25	95.1	341	3	US-09-014-116-48	Sequence 48, Appl 1
2	3.25	95.1	9401	2	US-08-432-693-1	Sequence 1, Appl 1
3	3.25	95.1	9416	3	US-08-811-566-19	Sequence 19, Appl 1
4	3.25	95.1	9595	3	US-09-014-416-4	Sequence 4, Appl 1
5	3.25	95.1	9595	3	US-09-014-416-6	Sequence 6, Appl 1
6	3.23	94.6	341	2	US-08-404-209-1	Sequence 1, Appl 1
7	3.23	94.6	341	3	US-08-854-531-4	Sequence 4, Appl 1
8	3.23	94.6	341	3	US-08-439-996-1	Sequence 1, Appl 1
9	3.23	94.6	341	3	US-09-014-416-47	Sequence 47, Appl 1
10	3.23	94.6	341	4	US-08-869-380-4	Sequence 4, Appl 1
11	3.23	94.6	341	5	PCT-US95-13552-4	Sequence 4, Appl 1
12	3.23	94.6	342	3	US-08-474-7008-39	Sequence 39, Appl 1
13	3.23	94.6	350	2	US-07-863-622-1	Sequence 1, Appl 1
14	3.23	94.6	350	5	PCT-US93-02266-1	Sequence 1, Appl 1
15	3.23	94.6	686	4	US-08-968-1218-37	Sequence 37, Appl 1
16	3.23	94.6	686	4	US-08-397-2208-25	Sequence 25, Appl 1
17	3.23	94.6	780	3	US-08-474-7008-45	Sequence 45, Appl 1
18	3.23	94.6	923	4	US-08-869-380-1	Sequence 1, Appl 1
19	3.23	94.6	923	5	PCT-US95-13552-14	Sequence 14, Appl 1
20	3.23	94.6	9401	1	US-07-910-760-9	Sequence 9, Appl 1
21	3.23	94.6	9401	1	US-08-440-519-9	Sequence 9, Appl 1
22	3.23	94.6	9401	5	PCT-US91-02225-9	Sequence 9, Appl 1
23	3.22	94.2	341	3	US-09-014-416-49	Sequence 49, Appl 1
24	3.22	94.2	9599	3	US-09-014-416-2	Sequence 2, Appl 1
25	3.22	94.2	9646	3	US-08-811-566-1	Sequence 1, Appl 1
26	3.22	94.2	12980	3	US-08-811-566-5	Sequence 5, Appl 1
27	3.15	92.2	14999	1	US-08-324-977-3	Sequence 3, Appl 1

28	315.4	92.2	149.9	2	US-08-384-616-3	Sequence 3, Appl1
29	315.4	92.2	149.9	2	US-08-304-666A-3	Sequence 3, Appl1
30	315.4	92.2	149.9	4	US-09-315-850-3	Sequence 3, Appl1
31	315.4	92.2	9416	1	US-08-324-977-1	Sequence 1, Appl1
32	315.4	92.2	9416	2	US-08-384-616-1	Sequence 1, Appl1
33	315.4	92.2	9416	2	US-08-304-668A-1	Sequence 1, Appl1
34	315.4	92.2	9416	2	US-09-315-850-1	Sequence 1, Appl1
35	311.4	91.1	334	2	US-08-470-426B-15	Sequence 15, Appl1
36	311.4	91.1	1863	2	US-08-470-426B-14	Sequence 14, Appl1
37	306.6	89.6	334	2	US-08-470-426B-1	Sequence 1, Appl1
38	306.6	89.6	1863	3	US-08-470-426B-13	Sequence 13, Appl1
39	301.6	88.2	9185	3	US-08-444-818-12	Sequence 12, Appl1
40	301.6	88.2	9185	3	US-08-444-818-13	Sequence 13, Appl1
41	301.6	88.2	9379	4	US-09-388-874-1	Sequence 1, Appl1
42	301.6	88.2	9379	4	US-09-388-874-1	Sequence 1, Appl1
43	293.6	85.8	9589	1	US-07-925-695-2	Sequence 2, Appl1
44	293.6	85.8	9589	1	US-07-925-695-2	Sequence 2, Appl1
45	292.8	85.6	9511	1	US-07-925-695-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1

Sequence 48, Application US/09014416
Patent No. 6153421

GENERAL INFORMATION:
APPLICANT: Yanagii, Masayuki

APPLICANT: Bukh, Jens
APPLICANT: Emerson, Susanne U.

APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND

FILE REFERENCE: 20264276

CURRENT FILING DATE: 1998-01-27

EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 65

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48

```

LENGTH: 341
TYPE: DNA

US-09-014-416-48

Query Match	95.1%;	Score 325.2;	DB 3;	Length 341;
Best Global Similarity	98.8%;	Pred. No. 4.6e-99;		
Matches 338;	Conservative 0;	Mismatches 3;	Indels 1;	Gaps 1;

[illegible]

Db 300 ggtgcttcgagtcgcccgaggagcttcgtagaccgtgcacc 341

RESULT 2

US-08-432-693-1
Sequence 1, Application US/08432693

GENERAL INFORMATION:

APPLICANT: Su, Michael
TITLE OF INVENTION: METHODS AND HOST CELLS FOR ASSAYING

TITLE OF INVENTION: EXOGENOUS AND ENDOGENOUS PROTEASE ACTIVITY
NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave

STREET: 1251 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: United States of America

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,693

FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr, James F

REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: VPI/95-01

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9090

TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 9401 base pairs

TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cDNA

HYPOTHETICAL: NO
ANTI-SENSE: NO

FEATURE:
NAME/KEY: mat_peptide

LOCATION: 3420..5312
OTHER INFORMATION: /product= "NS3 protease"

FEATURE:
NAME/KEY: mat_peptide

LOCATION: 5313..5474
OTHER INFORMATION: /product= "NS4A"

FEATURE:
NAME/KEY: mat_peptide

LOCATION: 5475..5552
OTHER INFORMATION: /product= "truncated NS4B"

US-08-432-693-1

Query Match 95.1%; Score 325.2; DB 2; Length 9401;
Best Local Similarity 98.8%; Pred. No. 1.8e-98;

Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Db 1 gccagcccccctgatggggcgacacatcacatagatacactccctgtgaggaactactg 60

Db 1 gccagcccccctgatggggcgacacatcacatagatacactccctgtgaggaactactg 60

Qy 1 gccagcccccctgatggggcgacacatcacatagatacactccctgtgaggaactactg 60

Qy 1 gccagcccccctgatggggcgacacatcacatagatacactccctgtgaggaactactg 60

Db 1 gccagcccccctgatggggcgacacatcacatagatacactccctgtgaggaactactg 60

Qy 1 gccagcccccctgatggggcgacacatcacatagatacactccctgtgaggaactactg 60

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 180

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 240

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 239

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 300

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 299

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 342

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Qy 121 ccccccctccggagagccatgctgcgcgaacccgctgagtagaccggaattgcacg 341

Db	121	CCCCCTCCCGGAGACCCATAGTGTCTCGGAAACCGGTAGTAACCGGAAATTGCCAG	180
Oy	181	gagacacggagttccttttttgatcaatccgcgcacatgccttgaagtttbgagctgcccc	240
Db	181	GACGACCGGGTCTCTTTCTTTGGAT-AAACCCGCTCAATGCTCTGAGATTTGGGGTCCCC	239
Oy	241	cgcgagactgctagccgagtagtgttggatctgcgaaagcctctgtgtactgctgatatg	300
Db	240	CGCAAGACTGCTATACCCGAGTAGTGTGGTGGGAAAGCGCTTGTGTACTGCTGATAG	299
Oy	301	gttgtctgcgagtgccccgcggaagcttcgtaagccgtgasc	342
Db	300	GCTGCTTCGAGTCCCGCCGGAGGTTCTCGTAGACCTGCACCC	341

RESULT 4
 US-09-014-416-4
 Sequence 4, Application US/09014416
 Patent No. 6153421
 GENERAL INFORMATION:
 APPLICANT: Yanag1, Masayuki
 APPLICANT: Bukh, Jens
 APPLICANT: Emerson, Susanne U.
 APPLICANT: Purcell, Robert H.
 TITLE OF INVENTION: CLOVED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
 TITLE OF INVENTION: USES THEREOF
 FILE REFERENCE: 20264276
 CURRENT APPLICATION NUMBER: US/09/014,416
 CURRENT FILING DATE: 1998-01-27
 EARLIER APPLICATION NUMBER: US 60/053,062
 EARLIER FILING DATE: 1997-07-18
 NUMBER OF SEQ ID NOS: 65
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 4
 LENGTH: 9595
 TYPE: DNA
 ORGANISM: Hepatitis C virus
 US-09-014-416-4

[illegible]

```

Patent No. 6153421
GENERAL INFORMATION:
APPLICANT: Yanaql, Masayuki
APPLICANT: Buhn, Jens
APPLICANT: Emerson, Susanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 20264276
CURRENT APPLICATION NUMBER: US/09/014,416
CURRENT FILING DATE: 1998-01-27
EARLIER APPLICATION NUMBER: US 60/053,062
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 65
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 9599
TYPE: DNA
ORGANISM: Hepatitis C virus
US-09-014-416-6

Query: Match          95.1%; Score 325.2; DB 3; Length 9599;
Best: Local Similarity 98.8%; Pred. No. 1,8e-98;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps

```

Query/Match	95.1%	Score 325.2	DB 3	Length 9599
Best/Local Similarity	98.8%	Pred. 0.1.8e-98		
Matches 338	Conservative 0	Mismatches 3	Indels 1	Gaps 1
QY	11	gccaagccccctgatalgaggcgagacatccaccataagatcatccctctgtaggaaactactg	60	
Db	1	gccaagccccctgatalgaggcgagacatccaccataagatcatccctctgtaggaaactactg	60	
QY	61	tcttcacgcagaaagcgtctctagccatgagcgttatgatatgagtgtctgtcagcgtcccaagcc	120	
Db	61	tcttcacgcagaaagcgtctctagccatgagcgttatgatatgagtgtctgtcagcgtcccaagcc	120	
QY	121	ccccctcccccggagagagcgcatagtgtgtctcggaaccggtttgattatcacccggaaattggccag	180	
Db	121	ccccctcccccggagagagcgcatagtgtgtctcggaaccggtttgattatcacccggaaattggccag	180	
QY	241	gacgacccgggtcctctctctcttgatcaatcccgctctcaatgcgtctgtaaaatttgggcgttgcgcc	240	
Db	241	gacgacccgggtcctctctctcttgatcaatcccgctctcaatgcgtctgtaaaatttgggcgttgcgcc	239	
QY	261	cgcgagactctgtagccggagtagtgttttgggtcgcgaaagcctttgtgttactgccttagatg	300	
Db	261	cgcgagactctgtagccggagtagtgttttgggtcgcgaaagcctttgtgttactgccttagatg	299	
QY	361	ggtgtctgtcgagtgcccccggagagctcgtgtaaacgctgacc	342	
Db	361	ggtgtctgtcgagtgcccccggagagctcgtgtaaacgctgacc	341	

RESULT 5
US-09-014-416-6
; Sequence 6, Application US/09014416

RESULT 6
 US-08-440-209-1
 Sequence 1, Application US/08440209
 Patent No. 5922857
 GENERAL INFORMATION:
 APPLICANT: Han, Jang H
 APPLICANT: Spaete, Richard R
 TITLE OF INVENTION: Methods and Compositions for Controlling
 TITLE OF INVENTION: Translation of HCV Proteins
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield, and Sacks P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,209
FILING DATE: 12-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,895
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: US 08/128,583
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Janluk, Anthony J
REGISTRATION NUMBER: 29,809
REFERENCE/DOCKET NUMBER: C0772/7004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 341 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: YES
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ns5hcv1
US-08-440-209-1

Query Match 94.6%; Score 323.6; DB 2; Length 341;
Best Local Similarity 96.5%; Pred. No. 1.6e-98;
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 gccagccccctgataggggcgagacactccacatagatactccctgtgaggaactactg 60
Db 1 gccacccccctgataggggcgagacactccacatagatactccctgtgaggaactactg 60
QY 61 tcttaccgagaagaagcgtctagccatgagcgttagatagatggtgtgtgagcctccagcc 120
Db 61 tcttaccgagaagaagcgtctagccatgagcgttagatagatggtgtgtgagcctccagcc 120
QY 121 cccccctccgggagagacatagtgcttcgagaaacggtgagacacggaattgccaag 180
Db 121 cccccctccgggagagacatagtgcttcgagaaacggtgagacacggaattgccaag 180
QY 181 gacgacgggttccttcttcttgatcaatcccgctcaatgcttgaggaatttgagcggtgc 240
Db 181 gacgacgggttccttcttcttgatcaatcccgctcaatgcttgaggaatttgagcggtgc 240
QY 241 cgcgagactgttagccgagtagtggtgtgtgctgagaaagccttggtgactgcctgataag 300
Db 241 cgcgagactgttagccgagtagtggtgtgtgctgagaaagccttggtgactgcctgataag 300
QY 301 ggtgctgcgagtgcccgagaggtctcgtlagaccgtgcacc 342
Db 301 ggtgctgcgagtgcccgagaggtctcgtlagaccgtgcacc 342

RESULT 7
US-08-854-531-4
Sequence 4, Application US/08854531
Patent No. 6025341
GENERAL INFORMATION:
APPLICANT: Wands, Jack
APPLICANT: Tokushige, Katsutoshi
APPLICANT: Makita, Takaji
TITLE OF INVENTION: CHIMERIC HEPATITIS B/HEPATITIS C VIRUS VACCINE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6025341r1s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia

STATE: PA.
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,531
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: APOL-0214
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 341 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-854-531-4

Query Match 94.6%; Score 323.6; DB 3; Length 341;
Best Local Similarity 96.5%; Pred. No. 1.6e-98;
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 gccagccccctgataggggcgagacactccacatagatactccctgtgaggaactactg 60
Db 1 gccacccccctgataggggcgagacactccacatagatactccctgtgaggaactactg 60
QY 61 tcttaccgagaagaagcgtctagccatgagcgttagatagatggtgtgtgagcctccagcc 120
Db 61 tcttaccgagaagaagcgtctagccatgagcgttagatagatggtgtgtgagcctccagcc 120
QY 121 cccccctccgggagagacatagtgcttcgagaaacggtgagacacggaattgccaag 180
Db 121 cccccctccgggagagacatagtgcttcgagaaacggtgagacacggaattgccaag 180
QY 181 gacgacgggttccttcttcttgatcaatcccgctcaatgcttgaggaatttgagcggtgc 240
Db 181 gacgacgggttccttcttcttgatcaatcccgctcaatgcttgaggaatttgagcggtgc 240
QY 241 cgcgagactgttagccgagtagtggtgtgtgctgagaaagccttggtgactgcctgataag 300
Db 241 cgcgagactgttagccgagtagtggtgtgtgctgagaaagccttggtgactgcctgataag 300
QY 301 ggtgctgcgagtgcccgagaggtctcgtlagaccgtgcacc 342
Db 301 ggtgctgcgagtgcccgagaggtctcgtlagaccgtgcacc 342

RESULT 8
US-08-439-996-1
Sequence 1, Application US/08439996
Patent No. 6057093
GENERAL INFORMATION:
APPLICANT: Han, Jang H
APPLICANT: Spaete, Richard R
TITLE OF INVENTION: Method and Compositions for Controlling
TITLE OF INVENTION: Translation of HCV proteins
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield, and Sacks P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA

```

APPLICANT Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 20264276
CURRENT APPLICATION NUMBER: US/09/014,416
CURRENT FILING DATE: 1998-01-27
EARLIER APPLICATION NUMBER: US 60/053,062
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 47
LENGTH: 341
TYPE: DNA
ORGANISM: Hepatitis C virus
US-09/014-416-47

Query Match 94.6%; Score 323.6; DB: 3; Length 341;
Best Local Similarity 96.5%; Pred. NO. 1.6e-98;
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1

CY 1 gccagcccccgcgataggggcgacacatccaccatagatcattccctgttaggaactacty 60
D 1 gccagcccccgcgattggggggcgacactccaccatagatcattccctgttaggaactacty 60
CY 61 tcttaacgcagaagaacgcgtctagccatagcgcttagatagatgctcgtgcagcctccagacc 120
D 61 tcttaacgcagaagaacgcgtctagccatagcgcttagatagatgctcgtgcagcctccagacc 120
CY 121 cccccctcccgaggagagccatagatgctctgcggaaaccggttagtaccacggaaattggccag 180
D 121 cccccctcccgaggagagccatagatgctctgcggaaaccggttagtaccacggaaattggccag 180
CY 131 gacgacccggtcctctctctcttgatcaatcccgctcaatgctcgttagagattggcgctgcccc 240
D 131 gacgacccggtcctctctctcttgatcaatcccgctcaatgctcgttagagattggcgcgctgcccc 240
CY 241 cgcgagactgctagcgcgagtagtgttgggtcgcggaaggcctgttgtagtgcctgtag 300
D 241 cgcgagactgctagcgcgagtagtgttgggtcgcggaaggcctgttgtagtgcctgtag 300
CY 301 ggtgcttcgagtgcccccgaggagctcgttagaccgctgacc 342
D 301 ggtgcttcgagtgcccccgaggagctcgttagaccgctgacc 342

RESULT 10
US-08-859-380-4
Sequence 4, Application US/08869380
Patent No. 6235888
GENERAL INFORMATION:
APPLICANT: Pachuk, Catherine J.
APPLICANT: Wands, Jack
APPLICANT: Wakita, Takaji
APPLICANT: Zurawski, Jr., Vincent R.
APPLICANT: Coney, Leslie R.
TITLE OF INVENTION: Hepatitis C Virus Vaccine
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
ADDRESSEE: No. 6235888ris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA.
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/869,380

```

Query Match	94.6%	Score 323.6	DB 4	Length 341
Best Local Similarity	98.5%	Pred. No. 1.6e-98		
Matches 337	Conservative	0	Mismatches 4	Indels 1
				Gaps 1

```

RESULT 11
PCT-US95-13552-4
: Sequence 4 Application PC/TUS9513552
:
: GENERAL INFORMATION:
:
: APPLICANT: Wands, Jack
: APPLICANT: Tokushige, Katsutoshi
: APPLICANT: Wakita, Takeji
: APPLICANT: Pachuk, Catherine J.
: APPLICANT: Zurawski, Jr., Vincent R.
: APPLICANT: Coney, Leslie R.
:
: TITLE OF INVENTION: HEPATITIS VIRUS VACCINES
:
: NUMBER OF SEQUENCES: 18
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
: ADDRESSEE: Norris
:
: STREET: One Liberty Place, 46th floor
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

Query Match	94.68;	Score 323.6;	DB 5;	Length 341;
Best Local Similarity	98.58;	Pred. No. 1.6e-98;		
Matches 337;	Conservative	0;	Mismatches 4;	Indels 1;
				Gaps 1

RESULT 12
 US-08-47008-39/c
 Sequence 39, Application US/08474700B
 Patent No. 6001990
 GENERAL INFORMATION:
 APPLICANT: Wands, Jack
 APPLICANT: Wakita, Takaji
 APPLICANT: Moradpour, Darius
 TITLE OF INVENTION: ANTISENSE INHIBITION OF HEPATITIS C
 TITLE OF INVENTION: VIRUS
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,700B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/240,382
FILING DATE: 10 May 1994
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00786/279001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-474-700B-39

Query Match 94.6%; Score 323.6; DB 3; Length 342;
Best Local Similarity 98.5%; Pred. No. 1.6e-98;
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 gccagcccccgtatgggggcgaacatccaccatagatccctcttgaggaactactg 60
DB 342 gccagcccccgtatgggggcgaacatccaccatagatccctcttgaggaactactg 283
QY 61 tcttcaagcagaagcgctcctagcctatgctatgctgctgctgctgctgctgctg 120
DB 282 tcttcaagcagaagcgctcctagcctatgctatgctgctgctgctgctgctgctg 120
QY 121 ccccccccgagagagcagatgctgctgagagcgctgagagcagagcagatgctgctg 180
DB 222 ccccccccgagagagcagatgctgctgagagcgctgagagcagagcagatgctgctg 163
QY 181 gacgacccggtccttcttctgataatcccgctcaatgctgagatltggcgctgccc 240
DB 162 gacgacccggtccttcttctgataatcccgctcaatgctgagatltggcgctgccc 104
QY 241 cgcgagagctctagcagagtagtctgctgctgagagagcctgctgctgctgctgctg 300
DB 103 cgcgagagctctagcagagtagtctgctgctgagagagcctgctgctgctgctgctg 44
QY 301 ggtgcttgcagtgcccccgagaggtctcgtagacgctgaccc 342
DB 43 ggtgcttgcagtgcccccgagaggtctcgtagacgctgaccc 2

RESULT 13
US-07-863-622-1
Sequence 1, Application US/07863622
Patent No. 5858650
GENERAL INFORMATION:
APPLICANT: CELEBUSKI, JOSEPH E.
TITLE OF INVENTION: METHODS FOR INACTIVATING NUCLEOTIDE
NUMBER OF INVENTION: SEQUENCES AND METAL CHELATES FOR USE THEREIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
STREET: ABBOTT PARK ROAD
CITY: ABBOTT PARK

ADDRESSEE: ABBOTT LABORATORIES
STREET: D377/APED, ONE ABBOTT PARK ROAD.
CITY: ABBOTT PARK
STATE: ILLINOIS
COUNTRY: LAKE
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/863,622
FILING DATE: 19920403
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WONG, MEAN KHING
REGISTRATION NUMBER: 33561
REFERENCE/DOCKET NUMBER: 5159.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9556
TELEFAX: 708-938-3517
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-863-622-1

Query Match 94.6%; Score 323.6; DB 2; Length 350;
Best Local Similarity 98.5%; Pred. No. 1.6e-98;
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 gccagcccccgtatgggggcgaacatccaccatagatccctcttgaggaactactg 60
DB 342 gccagcccccgtatgggggcgaacatccaccatagatccctcttgaggaactactg 60
QY 61 tcttcaagcagaagcgctcctagcctatgctatgctgctgctgctgctgctgctgctg 120
DB 61 tcttcaagcagaagcgctcctagcctatgctatgctgctgctgctgctgctgctgctg 120
QY 121 ccccccccgagagagcagatgctgctgagagcgctgagagcagagcagatgctgctg 180
DB 121 ccccccccgagagagcagatgctgctgagagcgctgagagcagagcagatgctgctg 180
QY 181 gacgacccggtccttcttctgataatcccgctcaatgctgagatltggcgctgccc 240
DB 181 gacgacccggtccttcttctgataatcccgctcaatgctgagatltggcgctgccc 239
QY 241 cgcgagagctctagcagagtagtctgctgctgagagagcctgctgctgctgctgctg 300
DB 240 cgcgagagctctagcagagtagtctgctgctgagagagcctgctgctgctgctgctg 299
QY 301 ggtgcttgcagtgcccccgagaggtctcgtagacgctgaccc 342
DB 300 ggtgcttgcagtgcccccgagaggtctcgtagacgctgaccc 341

RESULT 14
PCT-US93-03266-1
Sequence 1, Application PC/TUS9303266
GENERAL INFORMATION:
APPLICANT: CELEBUSKI, JOSEPH E.
TITLE OF INVENTION: METHODS FOR INACTIVATING NUCLEOTIDE
NUMBER OF INVENTION: SEQUENCES AND METAL CHELATES FOR USE THEREIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: D377/APED, ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)